No. 833
$17.95

THE COMPLETE BROADCAST ANTENNA HANDBOOK— Design, Installation, Operation & Maintenance

By John E. Cunningham

TAB BOOKS
Blue Ridge Summit, Pa. 17214

THE COMPLETE BROADCAST ANTENNA HANDBOOK— Design, Installation, Operation & Maintenance

By John E. Cunningham

FIRST EDITION

FIRST PRINTING— JULY 1977

Printed in the United States
of America

Library of Congress Cataloging in Publication Data

Cunningham, John Edward, 1923-
The complete broadcast antenna handbook.

"No. 833."
Includes index.
1. Radio--Antennas. 2. Television--Antennas.
3. Radio--Transmitters and transmission. 4. Television
--Transmitters and transmission. I. Title.
TK6565.A6C86 621.3841'35 77-8520
ISBN 0-8306-7833-6

Preface

The one area of broadcast electronics that seems to be shrouded in mystery is the antenna-and-feeder system. The reason is that most engineers who find their way into broadcasting become involved with circuits and circuit theory much more than with field theory. The engineer's field theory, like a muscle, becomes weak with disuse. The engineer could, if he had the time, brush up on antenna and field theory so that he could handle his antenna problems more efficiently. In the past, however, this has not been an easy task.

Most of the texts on antennas have been either highly mathematical, or else too superficial to be of any value. Furthermore, the type of mathematics used in antenna work tends to be unfamiliar to one who doesn't use it on a regular basis. This means that to brush up on antenna theory the engineer would first have to brush up on vector mathematics, and the time required is rarely available.

In working with broadcast engineers for over 30 years, I have found that there are three factors that cause problems in studying antennas:

1. Most antenna texts present a great deal more information than is needed to enable one to operate and maintain an antenna system. The process of

culling out the unessential is difficult, and there is a tendency to give up.

2. The mathematical operations involved in calculating impedances and field intensities are not particularly difficult, but they are very tedious. This has been a serious obstacle in the past, but with the advent of the pocket electronic calculator, most of the tedious operations are eliminated.
3. The engineer is apt to confuse the unfamiliar with the difficult. This is probably the most serious obstacle. If one has a preconception that a particular field of study is difficult, he will manage to make it difficult. An English author of a most readable book on calculus once introduced the subject with the adage "What one fool can do, another can." This adage applies equally well to antennas. All that the average broadcast engineer needs to know about antennas can be mastered with a little persistence.

The book can be thought of as consisting of three parts. Chapters 1 through 4 review the basic principles that underlie all antenna and transmission-line operation. Concepts that most frequently cause trouble are reviewed in more detail. Chapters 5 through 16 deal with standard broadcast antennas. The standard broadcast antenna is such that the engineer must be concerned with all of the details of the system; therefore the treatment is quite detailed. Chapters 17 through 19 deal with FM and TV antenna systems. The approach here is completely different, because the FM or TV antenna is supplied as a manufactured component and most of the system is located at the top of a tall tower, where the engineer can't even gain access to it. In this case, he needs to know enough of the basic principles to understand manufacturer's specifications and interpret the few measurements that he can make. The remaining chapters are devoted to subjects that all antennas have in common.

I would like to acknowledge the contribution that so many of my associates have made to my understanding of antennas. To the late Dan and Bill Hutton, John Battison, Palmer Greer, Don Pauley, and Chris Payne; to George Bartlett, of the National Association of Broadcasters who has done much to spread the knowledge of antenna theory and practice by supporting many seminars on the subject; and especially to Carl E. Smith, who has shared unstintingly his unending knowledge of the subject. Last, but far from least, is my gratitude to Grace, whose encouragement and inspiration made this work possible.

John E. Cunningham

Contents

Chapter 1
Basic Principles

All of the subjects covered in this chapter are very elementary. Every broadcast engineer has studied them at least once. Nevertheless, a review that points out features that are directly applicable to transmission lines and antennas is in order. Every engineer or technician who has participated in technical bull sessions knows that differences of opinion involving the operation of a complicated piece of equipment such as a television transmitter are quickly resolved. Discussions about such very elementary concepts as charges, fields, and magnetism, however, often show that our understanding of very elementary principles is fuzzy, to say the least. This fuzziness regarding elementary principles is often the underlying reason why devices such as transmission lines and antennas are often poorly understood.

It is easy to see why elementary principles are not well understood. When we describe the operation of a complicated transmitter system, we describe it in terms of simpler units such as transistors, tubes, and resistors, knowing that our audience understands these simpler building blocks. When we get to something very basic, such as an electric charge, there are no component parts on which to base our understanding. We base our understanding on observations of experiments, rather than physical reasoning. Knowledge of this type is disseminated in terms of what are really mathematical fictions, such as fields and lines of force. If a civilization in outer space mastered electromagnetic radiation and came to earth, we would probably find that their basic concepts were much different from ours.

These concepts or mathematical fictions are important in that they are the only way we have of talking or writing about the subjects. They can, however, cause a great deal of confusion if the ground rules are not used properly. A case in point involves the speed of propagation of electric charges in conductors.

Every schoolboy knows that electricity travels at the speed of light, which is 186,000 miles, or 300,000,000 meters, per second. It is also common knowledge that the charge carrier in conductors is the electron. It isn't unusual, therefore, to find people visualizing current in a conductor as consisting of a stream of electrons traveling through the wire at the speed of light. This idea vanishes when we apply the principles of physics to the problem.

Physicists say that a current of one ampere corresponds to a flow of 6.4×10^{18} electrons per second. This is a goodly number of electrons, so at first glance this figure seems to support the earlier idea. However, though we know how many electrons pass a point in a second, we *need* to know how many are passing together before we can determine the speed of individual electrons. We see this when we note that cars traveling four abreast will only have to travel at a quarter of the speed of cars in single file to have the same number of cars pass a point in a given time.

It requires an unimaginable number of electrons passing a point to produce an ampere of current, but there is also an unimaginable number of electrons available in a conductor. If we accept the physicists' figure of about 10^{23} free electrons per cubic centimeter of copper, we can calculate that the speed of electrons in a No. 12 wire carrying one ampere is about 0.08 in. per second—a far cry from the speed of light.

The above figure is based on steady, direct current. In antennas and transmission lines, we are interested in alternating currents with frequencies of one-half to several hundred megahertz. At 1 MHz the signal changes direction every half microsecond. At the slow rate at which electrons move, they barely move at all before they change direction.

If the electrons in a conductor actually move so slowly, what is all of this about electricity moving at the speed of light? The fact is, a current-carrying wire is analogous to a hollow pipe filled with marbles. The instant a marble is pushed into one end of the pipe, however long the pipe, a different marble pops out of the other end. The pushing effect travels through the pipe at a fantastic speed even though the speed of the individual marbles is quite slow. In an electric circuit, when a charge is introduced into one end, the effect is felt at the other end almost instantaneously, as if the charge itself traveled at the speed of light.

Thus, although electron flow is a valid and useful concept in vacuum tubes, it hardly makes any difference in antennas whether we think of electrons, or simply of charges, without defining the charge carrier.

CURRENT CONVENTIONS

One of the more controversial subjects in electronics is the question of what convention should be adopted for the direction of current flow. For many years it was almost universally agreed that a current flowed from the positive pole of a battery, through the external circuit, back to the negative pole. The convention was used long after it was well known that the electrons which actually carry the charge flow in the opposite direction. With the advent of the vacuum tube, it became advantageous to consider the flow of electrons from the cathode to the plate as being the plate current, and in many texts, particularly those below the engineering level, the negative-to-positive convention was adopted. This made the explanation of vacuum-tube operation easy, but it means that the direction of the drop of potential in a circuit is considered the opposite of the direction of current flow. This is almost like considering water as flowing against the direction in which pressure is exerted.

There will always be some inconsistency in application, regardless of what convention is adopted. In most of this book, dealing as it does with high frequencies, we will have little occasion to concern ourselves with current direction, but in some of the explanations, we will have to consider the flow of charges. Since it is more convenient to consider a positive cause as producing a positive effect, we will consider current to consist of positive charges flowing from the point of higher or more positive potential. This will undoubtedly offend some readers at first, but the concept is easy to apply when one becomes accustomed to it.

CHARGES AND FIELDS

An earlier section said that current is a flow of electric charges, without defining what a charge is. This is where we get to a concept so fundamental that we have no other, more elementary, concepts that we can invoke to explain it. We know that like charges tend to repel each other, and unlike charges attract. *Charge* is the concept that we have invented to explain this repulsion or attraction. In Fig. 1-1 we have a metal ball suspended above the earth. When we close the switch in the circuit, current flows, charging the ball. Or we could say that the battery forced some of the electrons off the ball through the battery onto the earth. We also know that if we connect a conductor between the ball and earth, current will flow through the conductor until the ball is at the same potential as the earth. When the ball is charged by the battery, there is a potential difference of 100V between the ball and earth. After a conductor has been connected between the ball

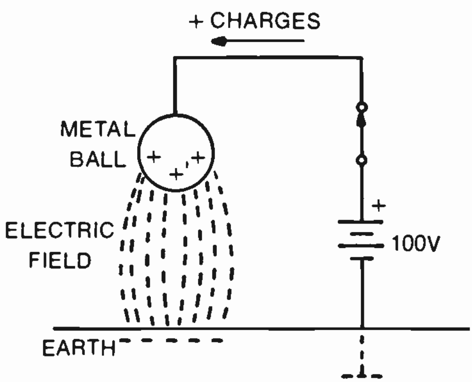


Fig. 1-1. Charging a metal ball.

and the earth for a short time, there is no voltage between them.

When the ball is charged, the electrons on the earth "know" that the ball is positively charged, and they are attracted to it. This attraction would take place even if there were absolutely nothing between the ball and earth. This *action at a distance* is repugnant to the average mind, so we say that there are *lines of force* between the positive charges on the ball and the negative charges on the earth. The lines of force are said to form an *electric field*.

Ether

The whole question of fields and lines of force can be very confusing. Back in 1865, long before the first radio signal was transmitted, James Clerk Maxwell, a Scottish physicist, theorized that light was actually an electromagnetic wave. His work implied that other electromagnetic waves might exist. About 20 years later a German physicist, Heinrich Hertz, actually demonstrated radio waves. To these early investigators, if light and electric and magnetic energy were propagated by a wavelike phenomenon, they must be waves in *something*. They didn't know just what this something might be, but they called it the *ether*.

Inasmuch as electric and magnetic fields were known to travel through a vacuum, they assumed that the ether permeated all space and matter. This concept was very useful for practical applications. Instead of speaking of the permeability or permittivity of free space, physicists could speak of the permeability and permittivity of the ether. It is much easier to attribute properties to *something*, even though we don't know just what the something is, than to attribute properties to free space, which, by definition, is nothing.

The concept of an ether was used by all of the early workers in radio, but scientists were troubled by the fact that no one had actually demonstrated the existence of the ether. They reasoned that if the ether did exist, either the earth moved through it, or it moved through the earth. Astronomical observations indicated that the earth must move through the ether. This meant that the velocity of light measured at the surface of the earth should be faster in one direction than in the opposite direction.

If a way could be found to measure the velocity of light in several directions on the surface of the earth, it should be possible to demonstrate the existence of the ether. In a now-famous series of experiments conducted between 1880 and 1890, two physicists—Michelson and Morley—measured the velocity of light to a high degree of accuracy. Much to their surprise they found that the velocity of light is the same in all directions along the surface of the earth.

The scientific community was faced with a dilemma: Some observations indicated that the earth moved through the ether, while others indicated that the ether moved with the earth—obviously a contradiction. After several futile attempts to explain the contradiction, the whole idea of an ether was dropped.

This was indeed unfortunate for the practical-minded engineer, who must now state that radio waves are propagated through empty space by an electromagnetic field—which is just another way of saying that we haven't the slightest idea of how radio waves are propagated.

This situation may be corrected in the next few years. An increasing amount of evidence is accumulating that the early investigators might have been right, and that there really is some sort of medium that carries radio waves, electric fields, and magnetic fields. In Europe two Nobel Prize winners, Dirac and De Broglie, have proposed that some sort of ether does exist. In this country Professor H. C. Dudley, of the University of Illinois, has written several papers that shed new light on the subject. Dudley points out that recent discoveries indicate that the whole universe is filled with a veritable sea of extremely small particles called *neutrinos*. He proposes that this neutrino is actually the medium that carries radio waves.

If Dudley's work proves to be correct—and there is an increasing amount of evidence that it is—the engineer will have a much clearer idea of what is actually going on in circuits and antennas. In the meantime, the reader should adopt whatever concept is most comfortable to him, with the consolation that, at present, the scientists don't actually know much more about it than he does.

Charge and Capacitance

Getting back to Fig. 1-1, if we were to repeat the experiment with a larger ball, or with the ball closer to the earth, we would find that more current would be required to charge the ball to 100V. This shows that voltage is not a good measure of how much charge we have. Actually the current in amperes that flows into the ball to charge it is the *rate of flow* of charge. The unit of measurement of charge is the *coulomb.* A current of *one ampere* means that charge is flowing at the rate of *one coulomb per second.*

If different arrangements similar to that of Fig. 1-1 will take on a different amount of charge for the same value of applied voltage, we need some measurement that will tell us how much charge each arrangement will take with a given value of applied voltage. We do have such a unit in the *farad.* The ability of a physical arrangement to acquire a charge when a voltage is applied to it is called its *capacitance*, the basic unit of which is the farad. The amount of charge in a capacitor is given by the equation

$$q = CV$$

where q is the charge in coulombs, C is the capacitance in farads, and V is the voltage across the capacitor in volts.

The farad, like so many basic units, is not of a very convenient size. In radio work we more commonly use microfarads or picofarads.

So far we have assumed that there was only air, which is electrically about the same as free space, between the ball of our experiment and the earth. If we were to fill this space with a material such as polystyrene, we would find that the ball took on more charge for the same value of applied voltage. As a matter of fact, it would take on just about twice as many coulombs of charge. We account for this by saying that the relative value of the permittivity, or the *dielectric constant*, of polystyrene is twice that of free space. The actual numerical value of dielectric constant depends on the unit system we are using. We will come back to this later.

Magnetic Field

By using analogies, we can also arrive at the concept of a magnetic field. If we were to pass a current through a coil of wire, we would find that it attracts pieces of magnetic material. We account for this attraction by saying that lines of magnetic force surround each current-carrying conductor.

We have no need in this book for the units used to describe magnetic quantities, but perhaps the analogy with electric fields will be a little clearer if we state them briefly. We measure the ability of a current to produce magnetic effects in terms of the *magnetizing force H*.

In a straight conductor the magnetizing force is expressed in amperes per meter and is numerically equal to the current in the conductor divided by its length. If we coil up the wire so that the magnetic effects of the turns reinforce each other, we usually state the magnetizing force in *ampere-turns per meter*. As with the electric field, the strength of the magnetic field can be measured in terms of the density of the lines of force. The unit of measurement of flux density B is the *weber per square meter*.

Everyone knows that a magnetic field is stronger in ferromagnetic materials than in free space. To explain this, we have the simple equation

$$\mu = B/H$$

where μ is the permeability of the material through which the magnetic field passes. Here again, the numerical value depends on the unit system that we are using.

UNIT SYSTEMS

There are several different systems of units used to specify different physical quantities. Each of these has its advantages and disadvantages. For example, there are two *cgs* (centimeter–gram–second) unit systems. In the cgs electro*static*-unit system, the permittivity of free space is simply 1. This makes calculation of capacitance easy, but to keep the system consistent, the unit of voltage becomes the *statvolt* and the unit of current becomes the *statampere*. These are both oddball units that will not ring a bell with the average broadcast engineer.

In the cgs electro*magnetic*-unit system, the *permeability* of free space is 1. This simplifies the calculation of magnetic field and inductances but leads to the *abvolt* and *abampere* as units of voltage and current which are as unfamiliar to the average broadcaster as the electrostatic units.

The unit system that is most widely accepted today is the so-called *rationalized mks* (meter–kilogram–second) unit system. In this system all of the commonly used quantities are expressed in familiar units such as volts, amperes, and ohms. The price that we pay for this very convenient system of units is that the permittivity ϵ_0 and permeability μ_0 of free space take on cumbersome values:

$$\epsilon_0 = 8.85 \times 10^{-12} \text{ farad per meter}$$
$$\mu_0 = 1.26 \times 10^{-6} \text{ henry per meter}$$

These two properties of empty space enter into the equation for the velocity of propagation of radio waves, as we shall see later on.

ENERGY AND POWER

Although power is given consideration in all parts of a broadcast station, energy is the more fundamental concept. *Energy* is defined as the *capacity to do work*. We are not interested in a strict definition of work. As far as we are concerned, it is sufficient to say that *work* is the *capacity to move something against an opposing force*. The object of any broadcasting system is to move electric charges in a receiving antenna somewhere. To do this requires energy. The energy must be carried from the transmitter, through the feeder system to the antenna, and there it must be radiated through space in the desired directions.

The basic unit of energy is the *joule*. Many engineers have forgotten this because they find it more convenient to work with *power*, which is the *rate of change of energy*. A power of one *watt* means a rate of change of energy of one joule per second. Although power is usually more convenient for practical calculations, it is much easier to understand things like reactance and reflections on transmission lines if we think in terms of energy.

Average Power

Figure 1-2 shows a source, such as an AC generator or oscillator, connected to a load. The resistance of the load is 10 ohms, which is assumed to be purely resistive, with no reactive component. We have a voltage of 20V across the load, and according to Ohm's law, the current will be 2A. Therefore we know that the power in the load is 20V × 2A = 40W. This means that energy is flowing into the resistive load at a rate of 40 joules per second. Without saying it, we realize that this is the *average* power, which is the power that we usually talk about. Common sense tells us that energy isn't actually flowing into the resistor at a constant rate. It must be zero at the instants in the cycle when both voltage and current are zero, and it must be maximum when both voltage and current are maximum.

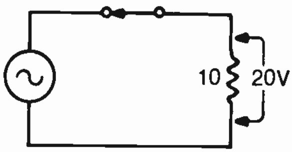


Fig. 1-2. A generator with a resistive load.

Instantaneous Power

It is helpful in understanding just how energy flows in a circuit to consider the *instantaneous* power, that is, the power at any instant of time. The instantaneous power in any circuit is equal to the product of voltage and current at some instant. Figure 1-3 shows a plot of the voltage and current and their product for the circuit of Fig. 1-2. During half of each cycle both voltage and current are positive, and during the other half they are both negative. Since the product of two negative numbers is positive, the power is positive at all times and varies at twice the frequency of the applied voltage. Thus the energy enters the resistor in pulses. In purely resistive circuits this pulsating nature of power and energy rarely concerns us, so we speak of the *average* power, which is the average value of the power wave in Fig. 1-3.

Energy in an Electrical Circuit

There are some other rather useful principles illustrated in the circuit of Fig. 1-2. Since the circuit is resistive, the

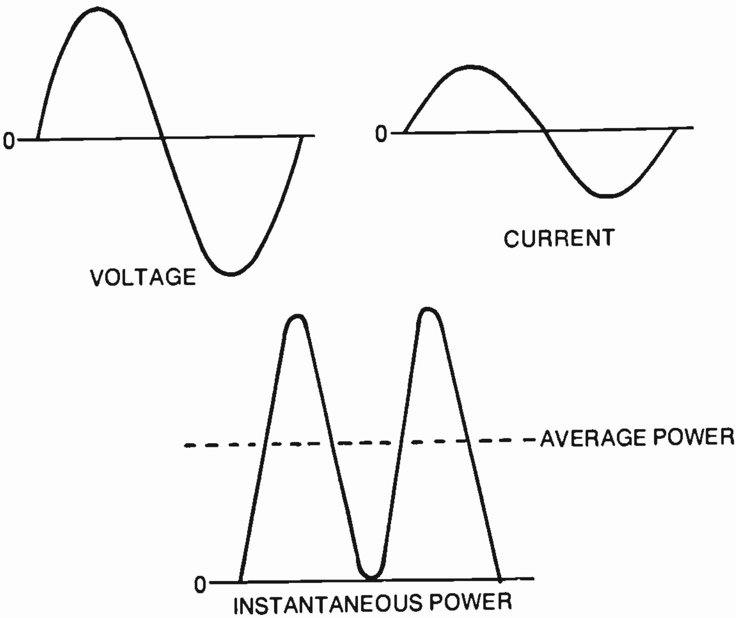


Fig. 1-3. Instantaneous and average power.

voltage and current are in phase at all times. This resistive nature of the load—or what is the same thing, the fact that the voltage and current are in phase—tells us that the energy is flowing in one direction and is not coming back. If the load is a resistor, the energy is converted into heat. If the resistive load represents the resistive component of some sort of motor, the electrical energy may be converted into mechanical energy. And if the resistance is seen at the terminals of an antenna, the energy is radiated. Although the theory of relativity shows that there is an interchange between matter and energy in some instances, as far as we are concerned the old law of conservation of energy still holds: Energy can neither be created nor destroyed; it is merely converted from one form to another.

In Fig. 1-4 we assume that a source is connected to a load that is purely capacitive and has no losses. In this case, the voltage and current are no longer in phase. The current leads the voltage by 90°. This is logical since, at the instant that the source is connected, there is no charge in the capacitor. Our earlier equation for the voltage across a capacitor can be written

$$V = q/C$$

This equation says there can be no voltage across a capacitor unless there is a charge in it. So in our circuit, as the current flows into the capacitor, charging it, a voltage builds up across it.

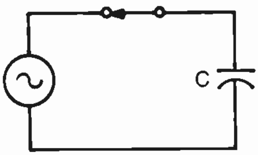


Fig. 1-4. A generator with a capacitive load.

We know that capacitors do not dissipate energy unless they have losses, so the average power in our circuit must be zero. But we also know that charge flows into and out of the capacitor and that this must represent some energy. The situation becomes clear when we consider the instantaneous power as we did in the resistive circuit. Figure 1-5 shows the voltage and current, and their product, the instantaneous power. The instantaneous power has twice the frequency of the applied voltage, but in this case it isn't positive all of the time. There are portions of the cycle where the voltage and current have opposite signs and their product is thus negative. *Negative power* is the rate at which energy flows *toward* the source (generator). The curve shows that energy enters the capacitor for a quarter of a cycle, then returns to the source during the next quarter-cycle.

The important point in this example is that when voltage and current are 90° out of phase, the net transfer of energy from the source is zero. Energy does in fact leave the source, and it is for a time stored in the electric field in the capacitor; but it is later returned to the source. This brings up the question of whether or not we can relate the amount of energy stored in the capacitor to the voltage that exists across it. We can do so by the equation

$$W = 1/2CV^2$$

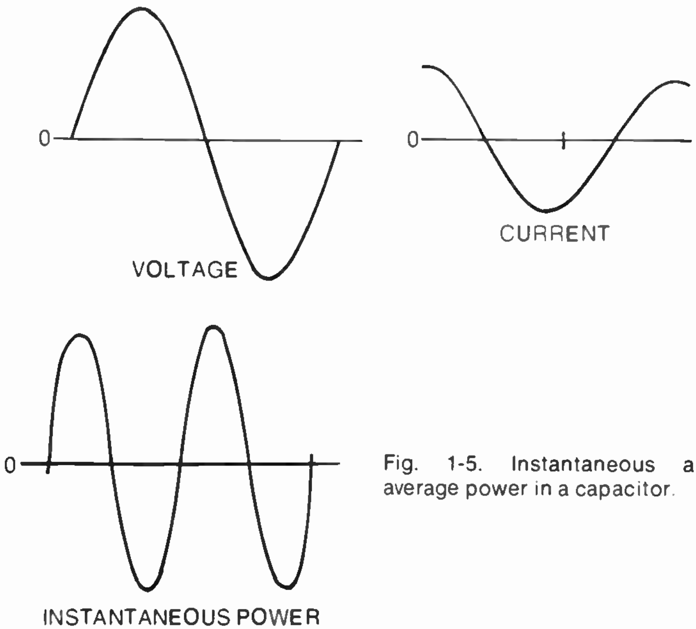


Fig. 1-5. Instantaneous and average power in a capacitor.

where W is the energy in joules, C is the capacitance in farads, and V is the voltage across the capacitor.

We know that if the load in our simple circuit were a pure inductance with no losses, the voltage would lead the current by 90°. It isn't difficult to construct an analogy to the case where the load is capacitive. We can see that energy must flow in and out of the inductor, but no average energy is taken from the source. As with the capacitor, we have an equation for the amount of energy stored in the magnetic field of an inductance. It is

$$W = 1/2LI^2$$

where W is the energy in joules, L is the inductance in henries, and I is the current in amperes.

In both the capacitance and inductance, the actual energy is stored in the associated electric or magnetic field. In fact, electric and magnetic fields are the only media that we know of in which we can store energy in its electrical form (When energy is stored in a battery, it is actually stored in the form of chemical energy.)

This leads us to a rather amazing, but logical, concept. Consider the circuit of Fig. 1-6, which consists simply of a source connected to a load some distance away. We know that the energy must get from the source to the load before it can do anything. We always assume, without wondering why, that the energy simply travels along the wires to the load, where it does whatever work the system was designed to do. We never stop to think that an ideal wire—one that has no inductance or capacitance —does not meet any of the requirements we have considered above for storing energy. But when we realize that each wire has inductance, with its incidental magnetic field, and that there is capacitance and an electric field between the wires, we have all the requirements for storing energy. Since we also know that in a capacitor or inductor the energy is actually stored in the associated field, this brings us to the rather startling conclusion that electrical energy is actually carried in the fields associated with wires, and not in the wires themselves! The wires merely serve to guide the energy to where we want it to go. This concept is hard to accept at first. Once accepted, it certainly makes radiation, whereby energy is propagated through space without the benefit of any wires at all, a lot easier to understand.

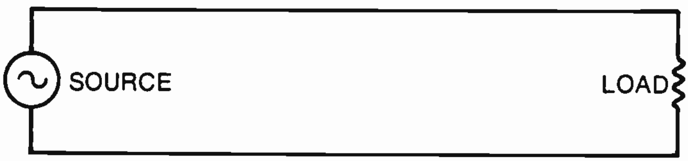


Fig. 1-6. Source with load at a distance.

SUPERPOSITION PRINCIPLE

One of the most useful concepts in all branches of physics, including antenna theory, is the superposition principle. The principle is very general and can be applied to any system, electrical or mechanical, wherein the elements of the system are *linear*. A linear element is one in which the response is directly proportional to the cause. Linear elements include resistances, inductances, and capacitances, but do not include such things as diodes. For our purposes the superposition principle can be stated:

In any system containing only sources of energy and linear bilateral elements—such as resistances, inductances, capacitances, transmission lines, and antennas—the total response with all sources active can be found by algebraically adding the responses that would be produced by each source acting separately.

This is quite a mouthful. All it means is that, in a system containing several sources, we can find out what is happening in any part of the system by finding out what would happen if each source acted alone, then algebraically combining the results.

Superposition can be applied to any physical situation in which the responses are linear. For example, by reflecting on the principle, we can state with confidence that a bullet fired horizontally from a gun will strike the earth at the same time that a similar bullet would strike the earth if it was merely dropped from the same height.

One place where *superposition* is very useful is in finding the field produced at some point by several antennas in a directional system. We simply find the field that would be produced by each antenna acting alone, then algebraically combine the fields to find the total resulting field.

Figure 1-7 shows the application of *superposition* in finding the current in a series circuit. Of course, this particular problem could be solved by a much simpler method, but this exercise shows some of the subtleties involved in applying superposition. To apply the principle, we first make one of the batteries idle (shorted) and find the current that would be produced by the other battery; then we reverse the process. Thus in Fig. 1-7B we replace battery B2 with a short circuit and find the current to be 3A, flowing counterclockwise around the circuit. In Fig. 1-7C we replace battery B1 with a short circuit and find the current to be 2A, flowing clockwise around the circuit. Since the two component currents are flowing in opposite directions, the net current actually flowing in the circuit is their difference: 1A, flowing counterclockwise around the circuit.

There are several points about Fig. 1-7 that should be clearly understood. First of all, the component currents do not actually flow in the circuit of Fig. 1-7A. They are the currents

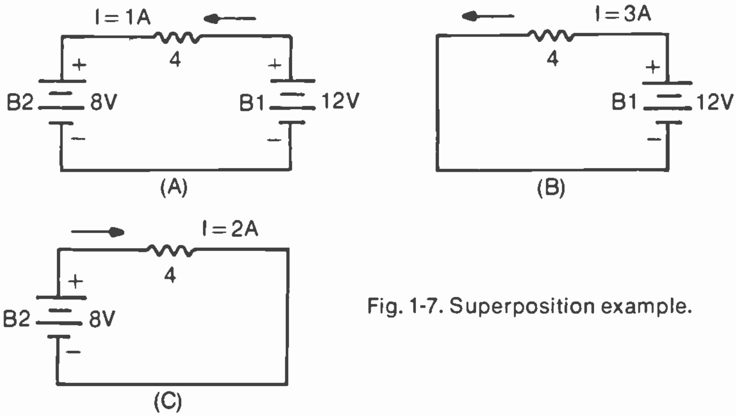


Fig. 1-7. Superposition example.

that *would* flow if each of the sources were acting alone, but the sources are *not* acting alone. The only current that actually flows in the circuit is the 1A flowing counterclockwise around the circuit. Battery B2 doesn't contribute anything to the actual current. In fact, the current is flowing "backward" through this battery in such a direction as to charge it. The point is that even though the component currents do not have a true physical existence, they are useful in finding the actual current. There are physical situations where applying the superposition principle provides not only a convenient way to get a numerical answer to a problem but also a better understanding of the physical principles involved.

We noted at the outset that superposition applies only when the response is linear. It is well to remember that *power is not a linear function of either the voltage or the current in a circuit*, and powers in a circuit or system cannot be found by simply adding the powers that each source would provide if acting alone.

Figure 1-8 shows another simple series circuit. The current I' due to battery B1 acting alone is 1A, and the current I'' due to battery B2 acting alone is 2A. The actual current in the circuit is 3A. Now, suppose that we tried to find the power in the 9-ohm load resistor by adding the power that would be supplied by each source acting alone. The power delivered to the load by each battery acting alone would be

$$P' = (I')^2 R = (1)^2 9 = 9\text{W}$$
$$P'' = (I'')^2 R = (2)^2 9 = 36\text{W}$$

giving a total of 45W. Actually we know that the current in the 9-ohm resistor is 3A, and the power is

$$PI^2 R = (3)^2 9 = 81W$$

Thus if we attempted to use superposition to add powers, the result would be in error. The reason for this is that $(I')^2 + (I'')^2$ is definitely not equal to $(I' + I'')^2$.

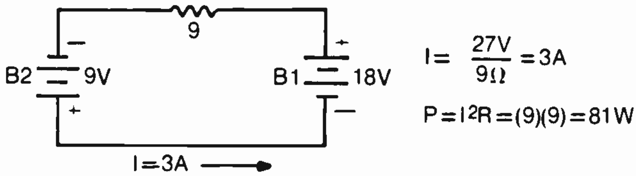


Fig. 1-8. Power in a simple series circuit.

WAVES, WAVELENGTH, AND DEGREES

When the term *wave* is used, one often thinks of waves in water. This is unfortunate because the waves we talk about in connection with radio are of a different nature. We use the word *wave* in two different senses. Radio waves are actually propagated by a wavelike action, but generally, when we use the word *wave*, we are speaking about a graphical representation of a physical phenomenon, not the action itself.

Sine Wave

The waveshape that concerns us most in antenna work is the sine wave. This wave is familiar to every broadcast engineer, but unless we keep our terminology clear, we get into some problems that will be very difficult to resolve. Figure 1-9 shows one cycle of the familiar sine wave. In this case, we will assume that it is the plot of voltage as a function of time. We find that the voltage smoothly increases to a positive peak, gently levels off, and smoothly decreases. It then does the same thing in a negative direction. The horizontal axis of our graph represents time. If the frequency of our wave was 1 MHz, the duration of one complete cycle, or *period* of the wave, would be 1 μsec.

This wave is also a plot of a trigonometric function that, happily, behaves in the same way as most of the voltages and currents that concern us. This means we can use trigonometric expressions to learn many things about our wave of voltage that would otherwise be hard to determine. The trigonometric function involved is the sine of an angle. If we vary the angle through 360°, the sine of the angle will vary just as the wave does in our illustration. This is shown by the set of numbers on the vertical axis of the graph. This is very convenient because it enables us to measure time periods in degrees. When doing this we should probably use the term *electrical degrees* to distinguish from degrees of arc, but common usage neglects this, and no trouble will be encountered if we keep the electrical concept clearly in mind.

To measure time in degrees is actually to measure time in fractions of the period of a wave. For example, if we were to state that another 1 MHz sine wave lags that of Fig. 1-9 by 90°, we would be saying that the new wave occurs 90° later in time. Since 90° is one-quarter of 360°, we could say that any point on the second wave occurs a quarter of a period, or 0.25 μsec, later than the corresponding point on the reference wave.

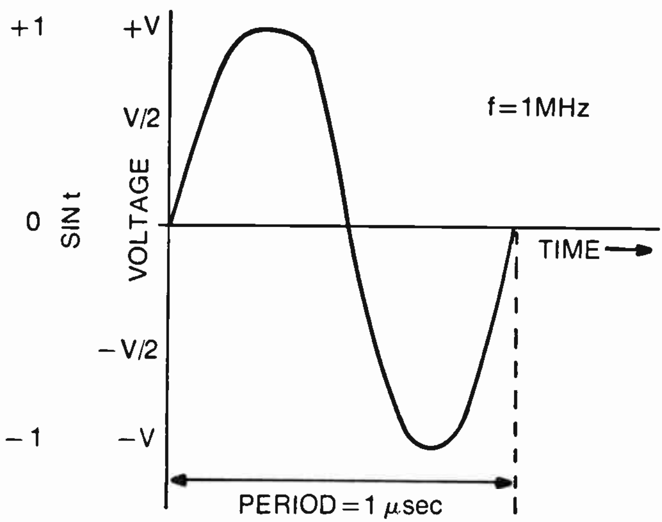


Fig. 1-9. Sine-wave plot, amplitude versus time.

The angle is usually called a phase angle. Note particularly that a phase angle (or phase shift) expressed in degrees can only be converted to time if we know the frequency (and hence the period) of the wave. Electrical degrees are not absolute, but relative, units of time. If the frequency in our example was 100 kHz instead of 1 MHz, 90° would correspond to a time period of 2.5 μsec rather than 0.25 μsec.

Wavelength

In broadcast work we are also interested in waves in space as well as waves in time. Suppose that the signal of Fig. 1-9 is radiated through space. Further suppose that we have a series of instruments spaced along the path that measure the signal strength at each point. Since the signal varies with time, it cannot have the same magnitude at all points in space at the same time. If at one instant we could stop time long enough to read all of the instruments and plot their indications as a function of distance along the path of propagation, the plot would be as shown in Fig. 1-10. The wave plotted here is identical in shape to the wave of Fig. 1-9, but it has a different meaning that should be clearly understood. Whereas in Fig. 1-9 we have amplitude versus time at some point in space, in Fig. 1-10 we have amplitude versus distance in space at a fixed instant of time.

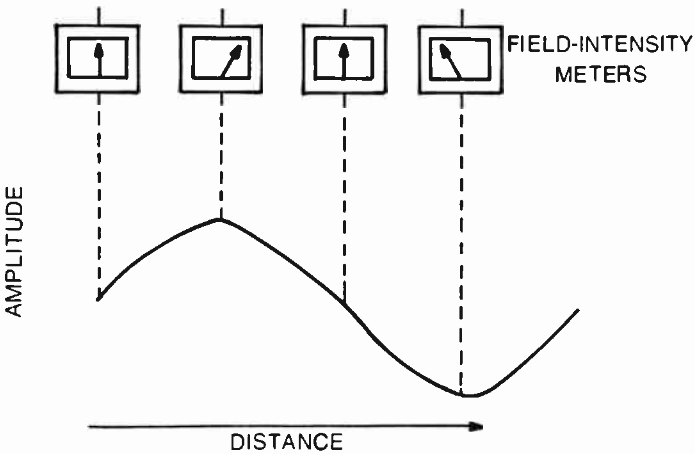


Fig. 1-10. Sine-wave plot, amplitude versus distance.

Since the wave of Fig. 1-10 is sinusoidal, we can express it in degrees, just as we did in Fig. 1-9. Here, however, a degree is an increment along the axis of the graph that represents *distance*, not time. Before we can correlate the degree with actual physical distance in space, we must take into consideration how fast the wave is moving through space. We can then tell how much distance is represented by one wavelength or a fraction of a wavelength, which may be expressed in degrees. The formula that we use to do this is

$$\lambda = \frac{c}{f}$$

where λ is the wavelength in meters, f is the frequency in hertz, and c is the velocity of propagation. In free space c is 300,000,000 meters per second. Thus our formula becomes

$$\lambda = \frac{300{,}000{,}000}{f \text{ in Hz}} \quad \text{or} \quad = \frac{300}{f \text{ in MHz}}$$

This formula shows that when we specify a distance in degrees, we are actually specifying the time it would take one wavelength of our signal to cover this distance. This is comparable to measuring the distance between two cities in hours when we know the speed at which we will travel between them.

There are two important facts to note about expressing distance in fractions of a wavelength or in degrees. First, the correlation with actual distance is only valid if we know the frequency. Second, the velocity of propagation enters into the relationship. As long as we are talking about propagation in free space, we can use the formulas given above; but when we get into a transmission line, where the velocity of propagation may be lower, we must make the necessary correction before we can correlate distance in degrees with actual physical distance.

In summary, the wave nature of radio signals makes it possible for us to measure either time or distance in degrees. In each case, frequency is the fundamental concept; we must know the frequency before we can do anything. The frequency is determined by whatever is generating the signal, usually a transmitter. When expressing distances in wavelengths or fractions of a wavelength, we must be sure of the velocity of propagation. In dealing with antennas, we encounter signals that can be represented by waves both as functions of time and functions of distance. Unless the distinction is kept clear in our minds, confusion will result.

The situation is further complicated by the fact that we must also consider radiation at various angles in space, which are also expressed in degrees. If the nature of each quantity expressed in degrees is kept clear, the situation isn't bad. It is only when we have a rather hazy idea of what we mean by *degrees* that confusion results.

PHASE LAG AND LEAD

The radiation properties of antenna arrays depend on the phase of the signals that are applied to various elements. When dealing with feeder systems, we must keep track of all of the phase shifts that are encountered, whether they are introduced by networks or by the time delay required for a signal to pass through a transmission line. To control these phase shifts, we use networks to introduce a desired amount of phase shift. These networks may either retard the phase of a signal and cause it to lag the input signal, or they may advance the phase and cause it to lead the input signal.

Phase Lag

The concept of a network that causes the output to lag the input is easy to accept. All we have to do is find something that will introduce a time delay. This will correspond to a phase lag, and knowing the frequency, we can find the number of degrees corresponding to any given time delay.

Phase Lead

The concept of a phase lead, or advance, isn't as easy to accept. It may seem that a network that causes the output to lead the input must move something forward in time. Putting it another way, it looks as though the network must have some way of "knowing" what the signal is going to be like in the future if the network's output is to look exactly as the input wave will at some time in the future. Of course, this isn't possible, and there is a better explanation of what happens.

Inasmuch as we are dealing with signals that do not vary much from one cycle to the next, even with modulation, we consider what is usually called the *steady-state* response that we get at the instant we apply power to the system. The situation is analogous to a platoon of troops following the orders of a drill sergeant. If the troops are at ease and not in formation and the drill sergeant suddenly barks out an unexpected command, the immediate response is nearly chaos while everyone gets in the proper position. Very soon afterward, however, the platoon is a smooth-functioning group that efficiently carries out all subsequent orders. The situation in an electric circuit is much the same. If the network is "at rest," with no charge in the capacitors and no current in the inductors, and a signal is suddenly applied, there follows a period during which the voltages and currents adjust themselves to the new environment. This is called the *transient* period. Very soon after this—the exact time depends on the *Q*-factor of the network—the network settles down and follows the dictates of the applied signal. It is this steady-state situation that concerns us.

Figure 1-11A shows an AC source connected through a switch and resistor to a capacitor. We will consider the voltage from the source to be the cause and the current in the circuit to be the effect. We know from elementary circuit theory that in a circuit containing both resistance and capacitance, the current leads the applied voltage—that is, the current reaches its maximum value before the voltage does. This looks like a clear case of the effect happening before the cause.

We can resolve this apparent difficulty by looking at the voltage and current in the circuit during the transient period. Let us assume that the switch in the circuit is closed at the instant when the applied voltage is at its maximum value. The voltage across the source, resistor, and capacitor, as well as the current in the circuit, are shown in Fig. 1-11B. At the instant the switch is closed, there is no charge in the capacitor, so it looks like a short circuit. The current will be maximum, but all of the voltage will appear across the series resistor. At this instant the current is in phase with the applied voltage; that is, the effect is occurring at the same time as the cause. From this time on, the voltage across the capacitor increases and the source voltage decreases.

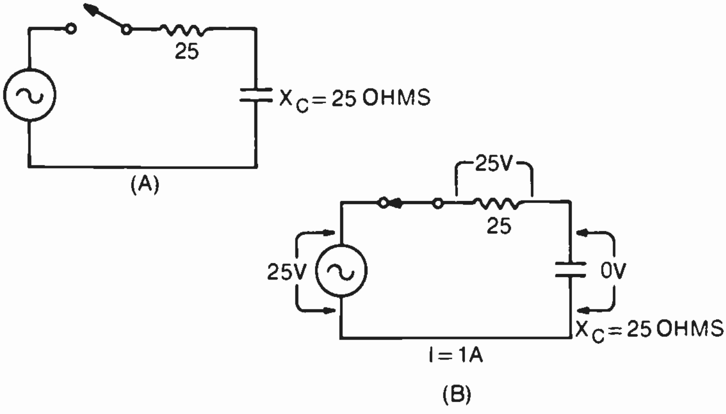


Fig. 1-11. Transient conditions in a circuit.

It is easy to see that the current in the circuit will stop flowing when the source voltage and the voltage across the capacitor are equal. After this the capacitor voltage will be higher than the source voltage, so the current will start to flow in the opposite direction. Thus the current, which is really the effect, reverses direction before the voltage, which is the cause. The phase angle between voltage and current is determined by the values of resistance and capacitance in the circuit. In this case, the phase angle is 45°. Since the applied signal is sinusoidal, once the circuit reaches the steady state the current will always lead the voltage by 45°.

VECTORS

A *vector* quantity is one that has both magnitude and direction. Quantities such as force and velocity, which have both magnitude and direction, are vector quantities. We can specify them in several different ways. For example, if the wind is blowing at 5 miles per hour from a direction which is 30° from north, we can specify the magnitude and velocity of the wind by a vector, which we write as 5/30°. The 5 indicates that the magnitude is 5 mph, and the 30° gives the direction. We can represent this vector graphically by drawing a line 5 units long at an angle of 30° from some reference line, which is usually, but not necessarily, the horizontal axis. The line itself is usually called a vector, and it is understood that the line is a graphical way of representing both the magnitude and direction of some quantity. The $5/30°$ vector is shown in Fig. 1-12.

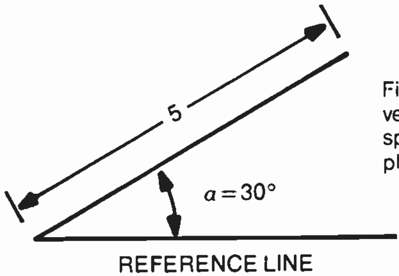


Fig. 1-12. Vectors and phasors. In vector notation a is an angle in space. In phasor notation a is a phase angle.

Vectors were originally used to specify physical quantities that had some direction in space. It was found, however, that a vector could be used to represent sinusoidal electrical quantities. When used in this application, the lines are more properly called *phasors*.

Use of Phasors and Vectors

There is a difference between a vector and a phasor: Vectors can easily be expanded to three dimensions, whereas phasors are restricted to two dimensions. Unfortunately, in an AM broadcast station the term *phasor* refers to a piece of equipment. This item was called a phasor long before the mathematical use of the word was coined. As a result, when the broadcast engineer hears the word *phasor*, he immediately identifies it with a piece of equipment. He has traditionally used the word *vector* to describe what the mathematicians call phasors, and this tradition will be respected throughout this book.

A vector can be used to represent a sinusoidal voltage or current, or an *impedance* (which is the ratio between them). For example, suppose that we wish to represent a current of 5A as leading the applied voltage, which we take for a reference, by 30°. Our vector diagram for this current is exactly the same as the vector diagram of Fig. 1-12. The only difference is that in the previous case the angle represented an angle in space whereas in the present case it represents a phase angle. We express the current in the same way as we expressed wind velocity: $5/30°$.

COMPLEX NUMBERS

A vector is just another way of writing a *complex number* (a number that can be resolved into two components at right angles to each other). To keep things on familiar ground, suppose that we have a circuit consisting of a 4-ohm resistance in series with a 3-ohm inductive reactance, as shown in Fig. 1-13A. We could write the impedance of the circuit as

$$\mathbf{Z} = 4 + j3 \text{ ohms}$$

Thus written, the impedance is said to be in *rectangular* form.

The use of the boldface symbol (**Z**) for impedance denotes that the vector is completely described, and not merely its absolute value given. To indicate the absolute value of a parameter, we use bars with the letter symbol, as in |Z|. In general, the use of boldface type and bars will not be necessary in this book, because the context of the problem or discussion will make clear which aspect of a quantity is of interest. In future chapters, this special symbology will be used only when required to prevent confusion.

We know from elementary AC theory that the magnitude of this impedance is

$$|Z| = \sqrt{R^2 + X^2} = \sqrt{16 + 9} = 5 \text{ ohms}$$

The phase angle between the applied voltage and the current is 36.87°. Now we can write the impedance in the form

$$\mathbf{Z} = 5 \angle 36.87^\circ$$

The above impedance can also be represented by the vector diagram of Fig. 1-13B. When written in this notation, the impedance is said to be written in *polar* form. As we will see, some mathematical operations are easier when complex numbers are written in polar form, and others are easier when the numbers are written in rectangular form. We convert from one form to the other by means of a right triangle and the two simple trigonometric expressions shown in Fig. 1-13C.

Adding and Subtracting Vector Quantities

Addition and subtraction of vector quantities is much easier when they are expressed in rectangular form. The procedure is as follows:

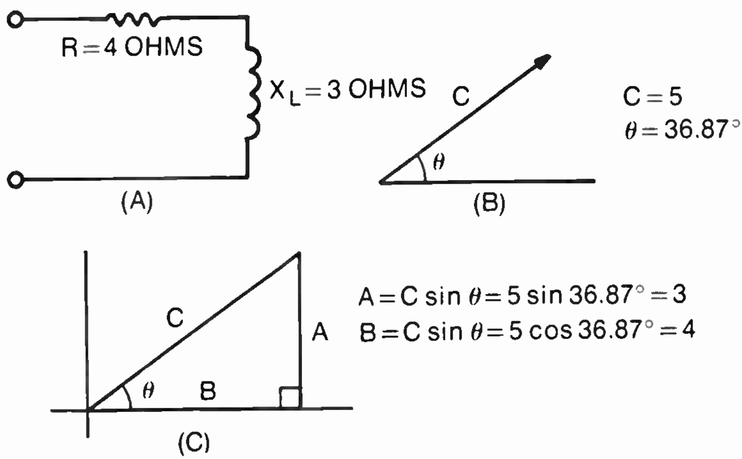


Fig. 1-13. Using vectors to represent electrical quantities.

1. Convert both numbers to rectangular form, using the method shown in Fig. 1-13C.
2. Add the real and "imaginary" parts separately.
3. Convert back to polar form if this form is required.

Figure 1-14 shows an example. Here we add $5 \angle 36.87°$

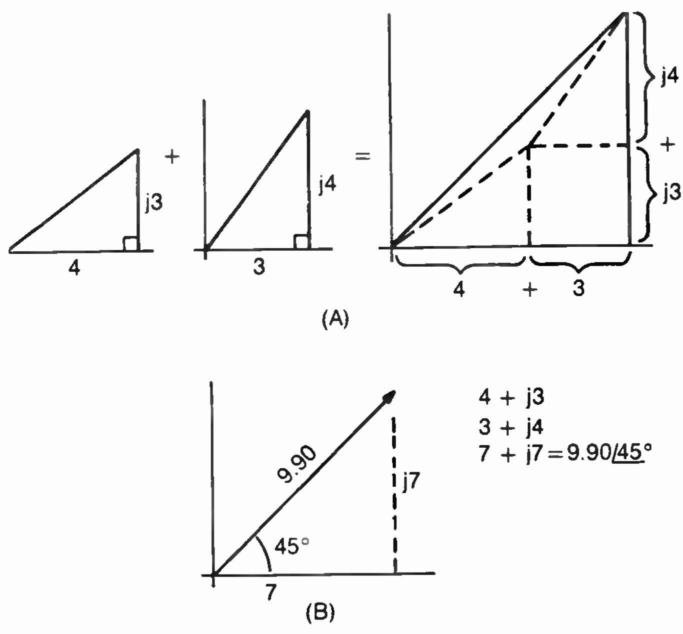


Fig. 1-14. Adding vectors.

and 5 $\angle 53.13°$. We use these odd angles to make the real and imaginary (j) numbers come out to whole numbers. The sum, when converted back to polar form, is 9.9 $\angle 45°$. In Fig. 1-15 we have an example of subtraction, which is simply the reverse of the addition of Fig. 1-14.

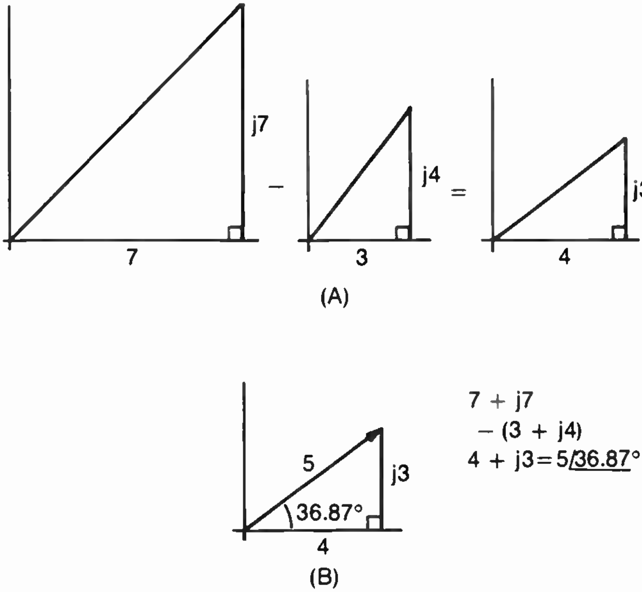


Fig. 1-15. Vector subtraction.

Multiplying and Dividing Vector Quantities

Two vectors can be multiplied together easily when they are expressed in polar form. All we have to do to find the product is to multiply the magnitudes and add the angles. For example, suppose that we wish to find the vector voltage across an impedance of 10 $\angle 15°$ when the current through the impedance is 5 $\angle 30°$ (i.e., 5 $\angle 30°$ amperes). From Ohm's law we know that we have to multiply current by impedance to find voltage. That is

$$\begin{aligned} \mathbf{V} &= \mathbf{IZ} \\ &= 5 \angle 30° \times 10 \angle 15° \\ &= (5 \times 10) \angle 30° + 15° = 50 \angle 45° \text{V} \end{aligned}$$

Division of vectors is just as easy. To reverse the problem just stated, suppose that we are given a voltage of $50 \angle 45°V$ and a current of $5 \angle 30°A$ and asked to find the impedance.

$$Z = V/I = \frac{50 \angle 45°}{5 \angle 30°} = \frac{50}{5} \angle 45 - 30 = 10 \angle 15° \text{ ohms}$$

Although vectors may be multiplied and divided in impedance problems, vector multiplication can *not* be used to find power.

In summary, we can express any complex number in either rectangular or polar form. In antenna work the polar form is useful for combining the field intensities from various antenna elements. In this book we will use whichever form tends to make clearer the technical points in question.

One does not actually have to go through the drudgery of performing the operations that we have presented. With an electronic calculator it is merely a matter of pressing keys. One should, however, have an understanding of the meaning of the quantities.

ANOTHER LOOK AT IMPEDANCE

The concept of impedance follows directly from Ohm's law for alternating currents and is familiar to every broadcast engineer. Nevertheless, when the concept is applied to such things as antennas and transmission lines, a great deal of confusion often results. For this reason we will briefly review the concept, with emphasis on some of its more subtle implications in antennas and transmission lines.

Fig. 1-16A shows a 10V source connected to two terminals on a box. To keep things simple for the moment, we will consider our source to be a battery. At the moment, we have absolutely no idea of what might be inside the box, but meters connected to the terminals tell us that when we apply 10V, the current will be 1A. We can then say that the impedance "looking into" the box is 10 ohms when the applied voltage is 10V.

If we are told that the box contains no nonlinear elements and no sources, we can assume that the impedance will also be 10 ohms for any other value of applied voltage. On this basis, we can assume that the equivalent of whatever happens to be in the box is a 10-ohm resistor, as shown in Fig. 1-16B. This does *not* mean that the box actually contains a 10-ohm resistor. It might, for example, contain two 20-ohm resistors connected in parallel (Fig. 1-16C).

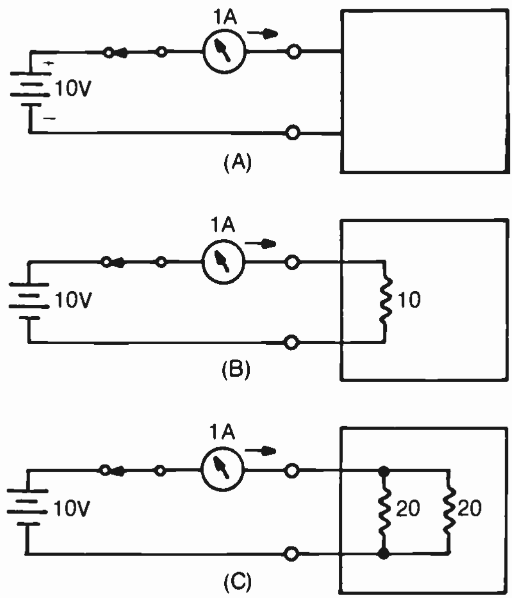


Fig. 1-16. Equivalent resistance.

So far the situation is very simple; there is no room for confusion. Things are a little more complicated in Fig. 1-17A, where we have an AC source and a method of measuring not only the current but also its phase angle, with the applied voltage as a reference. Our applied voltage is 10V, and the current in this case is 2A. We also find that the current lags the voltage by the now-familiar angle of −36.87°. (Since the current is lagging we prefix the phase angle with a minus sign).

The question is now "What is the equivalent circuit of whatever is in the box?" We can specify the ratio of the applied voltage to the current as an impedance. If we are only interested in the magnitude of the impedance, we can apply

Ohm's law just as we would in the DC case. We indicate the impedance that we find in this way as $|Z|$ to indicate that we know its magnitude but not its angle. Thus we can simply divide the magnitude of the voltage by the magnitude of the current to get the magnitude of the impedance.

$$|Z| = \frac{|V|}{|I|} = \frac{10}{2} = 5 \text{ ohms}$$

Of course, this doesn't tell us the whole story. We know that the current is not in phase with the applied voltage, so we know that there is some sort of reactive component in the box. To find both the magnitude and angle of the impedance, we can perform the same division in polar form.

$$\mathbf{Z} = \frac{\mathbf{V}}{\mathbf{I}} = \frac{10}{2 \angle -36.87^\circ} = 5 \angle 36.87 \text{ ohms}$$

This tells us that the box has an impedance that can be represented by the vector number $5 \angle 36.87^\circ$. The fact that the angle is positive indicates that whatever is in the box is inductive. We can derive an equivalent circuit for whatever is in the box by converting the impedance from polar to rectangular form, giving

$$\mathbf{Z} = R + jX_L = 4 + j3 \text{ ohms}$$

Thus we can say that the equivalent circuit of whatever is inside the box is a 4-ohm resistor in series with a 3-ohm inductive reactance (Fig. 1-17B). This doesn't mean that these elements are actually in the box. It just means that whatever actually is in the box will behave as these two elements at the frequency of interest. The box might, for example, contain a series-resonant circuit as in Fig. 1-17C.

Self-Impedance

So far we have been concerned with the impedances that are seen looking into the terminals of a box that contains only two terminals and no source. We call this impedance the *self-impedance* of the circuit inside our box. We could also call it the *driving-point* impedance seen at the terminals of the box

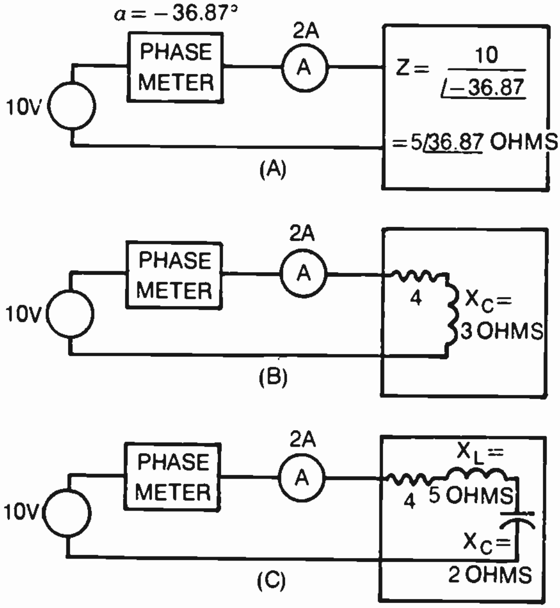


Fig. 1-17. Equivalent circuits containing inductance and capacitance.

As long as we have only one source of energy and one pair of terminals in a circuit, the self-impedance and driving-point impedance are really the same thing. Thus if we are talking about circuit elements such as resistors, coils, and capacitors, the driving-point impedance means the same thing as the self-impedance.

In Fig. 1-18A we have a completely different situation, and one that we will encounter very frequently in antenna work. Here the box has two sets of terminals. We will assume that the bottom terminals, 1′ and 2′, are connected together and grounded. Again, we have no idea whatever of what might be inside the box, except that it contains no sources or nonlinear elements. There are several measurements we could make that would enable us to draw an equivalent circuit for whatever happens to be inside the box.

We could, for example, apply a voltage to one pair of terminals while the other pair is open-circuited, as shown in Fig. 1-18B, and take the ratio of the applied voltage to the current. In Fig. 1-18B we see that if we apply 10V to the terminals at the left side of the box, a current of 1A will flow. If we divide the 10V by 1A, we get an impedance of 10 ohms. We call this the self-impedance between terminals 1 and 1′ and usually identify it as Z_{11}. Remember that the other terminals were open when we made this measurement. By a similar measurement we see in Fig. 1-18C that the self-impedance between terminals 2 and 2′, at the right (which we will call Z_{22}) side of the box is 5 ohms. These two measurements tell us what will happen when we energize either pair of terminals separately, but they give no information whatever on what connection may exist between the two sets of terminals or what will happen if we energize both sets of terminals at once.

Mutual Impedance

There is another measurement that will enable us to draw an equivalent circuit for whatever is in the box. We can connect our source to one pair of terminals and measure the voltage that appears across the other pair, as shown in Fig. 1-18D. The ratio of the voltage that appears between terminals 2 and 2′ to the current that is flowing in terminals 1 and 1′ is called the *mutual impedance* Z_{12} between the two sets of terminals. In Fig. 1-18D it is seen to be 2 ohms. It may be surprising at first, but as long as our box contains only linear circuit elements (resistances, inductances, and capacitances), it makes no difference to which terminals we apply the source. We could have connected the source between terminals 2 and 2′, and the voltmeter between terminals 1 and 1′, the mutual impedance would be the same in both cases.

$$Z_{12} = Z_{21}$$

This concept of mutual impedance between two sets of terminals should be clearly understood. It is probably responsible for more confusion about the behavior of antenna feeder systems than any other factor. Remember, the mutual impedance is the ratio of the voltage that appears across one pair of terminals to the current flowing in the other pair of terminals. The voltmeter used to measure this voltage must draw negligible current.

We are now in a position to draw an equivalent circuit for whatever happens to be in the box of Fig. 1-18A. There is a

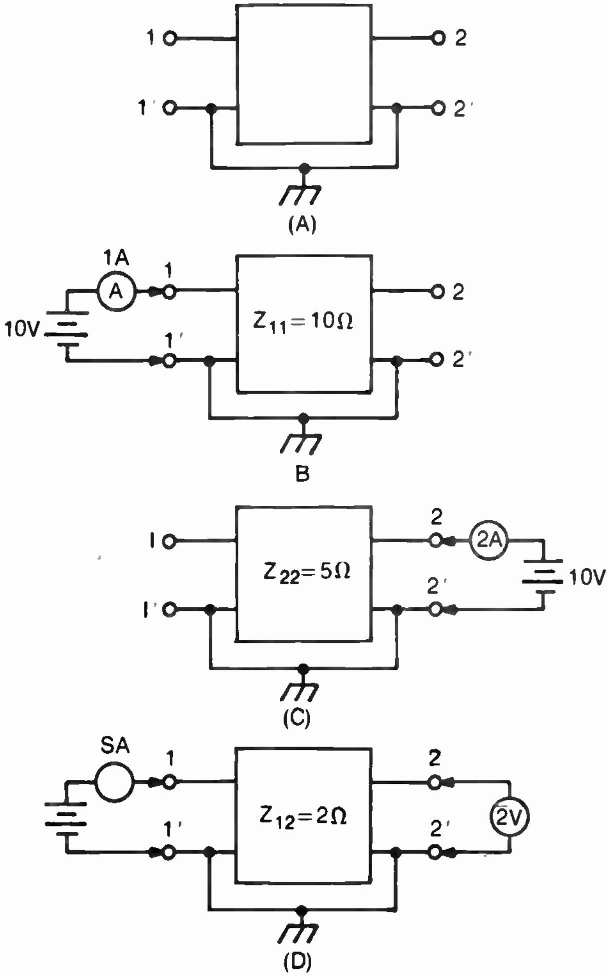


Fig. 1-18. Unknown circuits with two sets of terminals.

principle in circuit theory that any linear, bilateral circuit containing only linear passive components can be represented at one frequency by either a *T* or *pi network*. We will use the equivalent *T*-circuit because it will make the concepts involved clearer.

Figure 1-19A shows a *T*-network. We can use the measurements of Fig. 1-18 to find the values of the elements in this circuit that will make it behave exactly like the box. When terminals 2 and 2′ are open-circuited and current is flowing in terminals 1 and 1′, the output voltage will be simply $I_1 \times Z_b$. Therefore, Z_b is equal to the mutual impedance Z_{12}, in this case 2 ohms.

Next, we know that the impedance at terminals 1 and 1′ when the other terminals are open is the self-impedance between these terminals. We will call this impedance Z_{11}, and in our example it is 10 ohms. Now, if we are to see this impedance when we look into terminals 1 and 1′ with the other terminals open, then Z_a must be equal to $Z_{11} - Z_{12}$, in this case 8 ohms. Similarly, Z_c must be equal to $Z_{22} - Z_{12}$, where Z_{22} is the self-impedance between terminals 2 and 2′. We have now completely pinned down our equivalent circuit as shown in Fig. 1-19B.

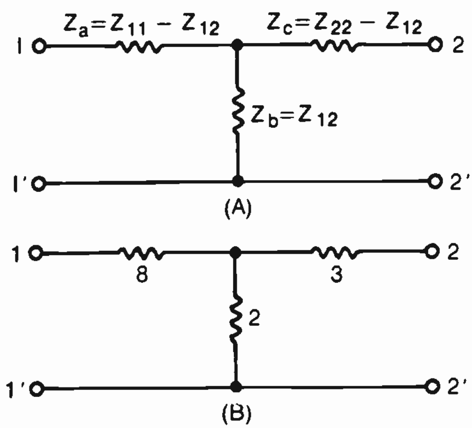


Fig. 1-19. Deriving an equivalent I-circuit.

Driving-Point Impedance

So far we have defined two kinds of impedance in connection with our circuit: the self-impedance of each pair of terminals, and the mutual impedance between the pairs. We know that when terminals 2 and 2′ are open, the *driving-point* impedance between terminals 1 and 1′—that is, the impedance seen looking into these terminals—will equal the self-impedance associated with them. The same is true of the other pair of terminals. But if there was something connected to both pairs of terminals at the same time, the driving-point impedance at one pair would probably not equal its self-impedance. We will not attempt to derive an expression for driving-point impedance into one pair of terminals that will hold regardless of what might be connected across the other pair.

Figure 1-20 shows our equivalent *T*-network with voltages applied to, and current flowing in, both sets of terminals. We can use Kirchhoff's voltage law to write an equation for the voltages around the loops in the figure. The equation for loop 1 is

$$\begin{aligned} V_1 &= I_1 (Z_{11} - Z_{12} + Z_{12}) - I_2 Z_{12} \\ &= I_1 Z_{11} - I_2 Z_{12} \end{aligned}$$

If we divide all of the terms of this equation by I_1, we get

$$\frac{V_1}{I_1} = Z_{11} - \frac{I_2}{I_1} Z_{12}$$

It is easy to see that this equation gives the ratio of the voltage between terminals 1 and 1′ to the current that will flow in them. Since the ratio of a voltage to a current is an impedance, we call this impedance the *driving-point* impedance between terminals 1 and 1′ and represent it by the symbol Z_1. Thus the equation for the driving-point impedance between terminals 1 and 1′ becomes

$$Z_1 = Z_{11} + \frac{I_2}{I_1} Z_{12}$$

This equation is very important and should be studied carefully. It shows that in a *T*-network, which is a good equivalent for many antenna circuits, the driving-point impedance depends on the currents flowing in both pairs of terminals. Remembering this simple equation will often remove confusion that results from interaction between the networks in antenna feeder systems.

We have now defined three different kinds of impedance associated with a network that has two sets of terminals. (1) The *self-impedance* (Z_{11} or Z_{22}) associated with a pair of terminals is the impedance seen looking into the terminals when nothing is connected to the other set of terminals. (2) The *mutual impedance* (Z_{12} or Z_{21}) is associated with *two pairs of terminals* and is the ratio of the voltage across one pair of terminals to the current flowing in the other pair. The magnitude of the mutual impedance depends on how the two pairs of terminals are connected together. In a more general sense, it depends on how energy gets from one pair of terminals to the other pair. (3) The *driving-point impedance* associated with a pair of terminals is the ratio of voltage to current at the terminals under certain conditions. The driving-point impedance depends not only on self-impedance but also on the currents flowing in the network.

To keep the mathematics comparatively simple, we have assumed that the impedances in our equivalent circuit were pure resistances. In general, this will not be true. Most of the impedances we encounter in antenna work will have reactive components. Thus the Zs in Fig. 1-20 would ordinarily have both magnitude and a phase angle. Furthermore, the current in loop 2 will often not be in phase with the current in loop 1. Thus we will have the phase angles of the currents to consider.

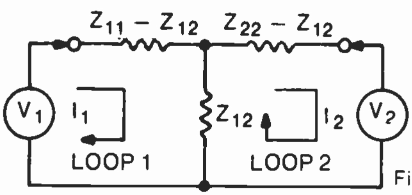


Fig. 1-20. Equivalent I-circuit with voltages and currents.

NEGATIVE RESISTANCE

A concept that rears its ugly head with annoying frequency in directional-antenna work is *negative resistance*. When a driving-point resistance found in an antenna feeder system turns out to be negative, it merely means that the current, and hence the power, is flowing *out* of the terminals rather than into them. This is illustrated in Fig. 1-21. Here we have the now-familiar *T*-network with two sources—one at each set of terminals. All of the voltages, polarities, and current directions are shown. We can now compute the driving point impedance at both sets of terminals; but in doing so, we should keep track of the polarity of voltages and the direction of current flow. We assume that a positive current is caused by a positive voltage. Thus the driving-point impedance Z_1 between terminals 1 and 1′ is

$$Z_1 = \frac{V_1}{I_1} = \frac{45V}{2.5A} = 18 \text{ ohms}$$

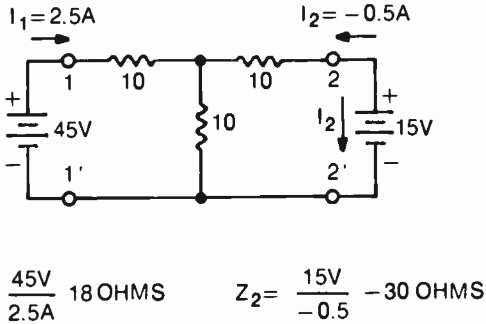


Fig. 1-21. Negative impedance.

Now we compute the driving-point impedance between terminals 2 and 2′:

$$Z_2 = \frac{V_2}{I_2} = \frac{15V}{-0.5A} = -30 \text{ ohms}$$

Because of the signs, we find that the impedance between these two points is a negative number. All this means is that power is flowing out of terminals 2 and 2′, as shown by the dashed line, instead of into them. A negative resistance means that energy is flowing opposite to the direction that it would flow in if the resistance were positive. The magnitude of the impedance is still simply the ratio of voltage to current at that point. Note that this is the only sense in which we shall use the concept of negative impedance. It is used in a different sense in connection with some semiconductor devices and oscillators, but that will not concern us. It is also important to note that only a driving point or mutual impedance can be negative; self-impedance is always positive.

The value of our equivalent circuit can be better appreciated by considering Fig. 1-22. Here we have a very large, albeit fictitious, box that contains two antennas instead of circuit elements. By using our equivalent *T*-network, we can reduce what would be a horrendous problem in field theory to a comparatively simple problem in circuit theory.

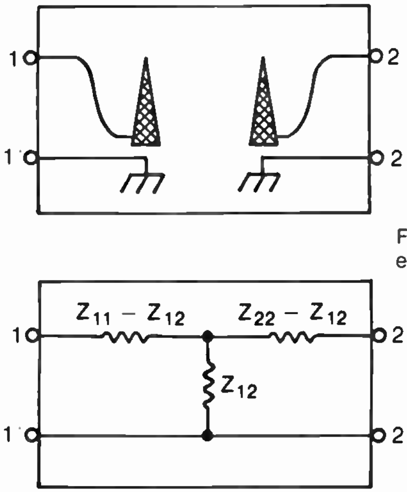


Fig. 1-22. Antenna system and equivalent circuit.

VECTORS AND POWER

Earlier, in applying Ohm's law, we found that we could multiply or divide vectors to find voltage, current, or impedance, all of which are vectors. The question naturally arises as to whether we can multiply voltage expressed as a vector by current expressed as a vector to find the power in a circuit. For example, if the voltage applied to a circuit was $10 \angle 0^\circ$ V and the current was $5 \angle -60^\circ$, could we multiply them to get power? The answer is no. If we were to multiply them we would get

$$10 \angle 0^\circ \times 5 \angle -60^\circ = 50 \angle -60^\circ$$

The fact that the product has an angle tells us that something is wrong. Power is simply the rate of flow of energy, and it doesn't have an angle. In other words, power is not a vector quantity.

The reason that this simple approach to finding power does not work is that the power moving past a point in a circuit is equal to the product of the voltage and the component of the current that is in phase with the voltage. Figure 1-23 shows a vector representation of our voltage and current. It can readily be seen that the component of the current that is in phase with the voltage is given by $I \cos \theta$, or 5×0.5, so the power in our circuit is given by

$$P = V(I \cos \theta) = 10 \times 5 \times 0.5 = 25\text{W}$$

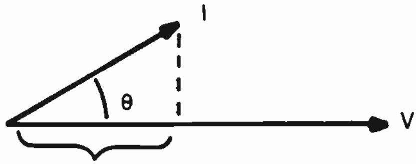


Fig. 1-23. Vector diagram of current of 5A displaced from voltage of 10V by 60°.

EQUIVALENT SOURCES, POWER TRANSFER, AND EFFICIENCY

So far we have had very good luck in drawing equivalent circuits for boxes containing only passive elements. We need an equally simple way of representing sources of energy, such as broadcast transmitters. This can be accomplished by a principle known as *Thevenin's theorem*, which is fully as useful as any of the principles we have investigated so far.

Thevenin's theorem states that any circuit that contains sources can be represented at one frequency by an ideal voltage source in series with an impedance. Of course, this equivalence only holds true over the operating range where everything is linear. We couldn't, for example, short-circuit the output of a transmitter and expect it to behave as a linear device—in fact, it probably wouldn't behave at all.

Over its normal operating range, however, we can expect a transmitter, signal generator, or almost any other source of power to look electrically like the equivalent circuit of Fig. 1-24. The voltage source V_T is an ideal constant-voltage source. It will produce the same output voltage regardless of what is connected to its terminals. The resistance R_S is the effective internal resistance.

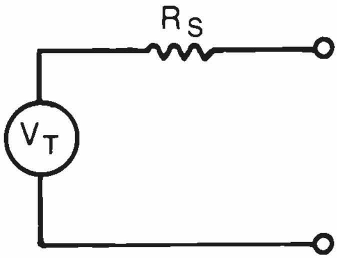


Fig. 1-24. Thevenin equivalent circuit.

Now that we have an equivalent circuit for a source, we can look at how load impedance will affect the amount of power we get out of the source. Figure 1-25A shows a source connected to a variable load resistance R_L. Figure 1-25B shows a plot of the power delivered to the load as a function of the ratio of the load resistance R_L to the equivalent source resistance R_S. It can be seen that the maximum power will be delivered to the load when $R_S = R_L$, that is, when the load resistance is equal to the internal resistance of the source.

On the surface, this looks like a very desirable situation, but a little deeper look will show that it probably isn't as attractive as it first seems. The same current flows through both the internal source resistance and the load resistance. When the two resistances are equal, just as much power is dissipated internally in the source as is delivered to the load. We get maximum power in the load, but the price we pay is an operating efficiency of only 50%. If the load applied to a

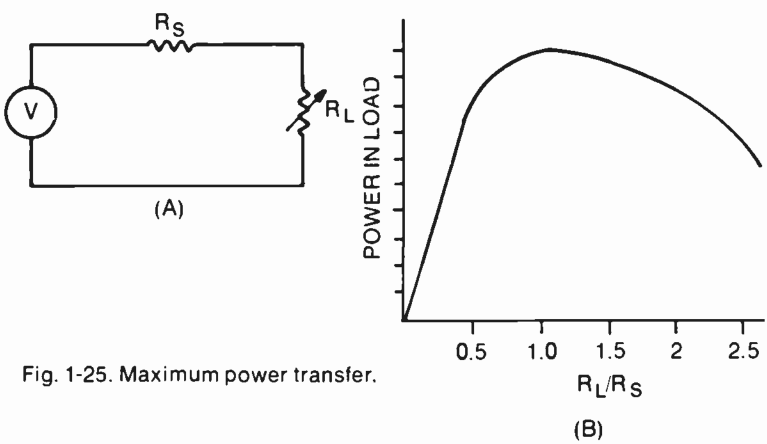


Fig. 1-25. Maximum power transfer.

broadcast transmitter were equal to its effective internal impedance, the efficiency of the final stage would only be 50%. Most transmitters are operated at much higher efficiencies.

In most cases, we don't know the internal impedance of a broadcast transmitter. The manufacturer normally specifies only the load impedance into which it is designed to work. Actually this is all we need to know to keep an antenna system operating properly. It is helpful, however, in trying to understand the operation of feeder systems, to get an idea of what impedance is seen looking into the transmitter output terminals. We can get a very rough idea of this from the specified load impedance and the operating efficiency of the final stage.

Suppose, for example, that we have a transmitter that is designed to work into a 50-ohm load at an efficiency of 70% (Fig. 1-26). The efficiency of this circuit expressed as a decimal, is given by

$$\text{eff} = \frac{R_L}{R_S + R_L}$$

Rearranging

$$R_S = R_L \left[\frac{1}{\text{eff}} - 1 \right]$$

Substituting numbers into this, we get

$$R_S = 50 \left[\frac{1}{0.7} - 1 \right] = 50[0.43] = 21.4 \text{ ohms}$$

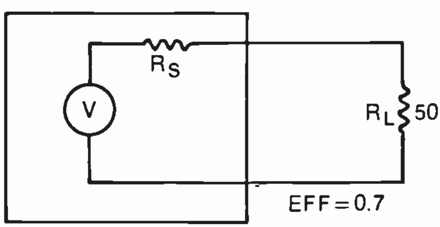


Fig. 1-26. Transmitter operating into 50Ω load with 70% efficiency.

Thus the internal impedance is about 21 ohms. This is only an approximation, but it shows that when a transmitter is operating at an efficiency of greater than 50%, its internal resistance is lower than the source impedance into which it operates. This means that a transmitter, in general, acts more like a constant-voltage source than a constant-current source. Looking at Fig. 1-26, we can see that if R_S was much lower than R_L, the *voltage* across the load wouldn't change much with small changes in load resistance. Also, if R_S was much larger than R_L, the load *current* wouldn't change much with small changes in load resistance.

CONDUCTANCE, SUSCEPTANCE, AND ADMITTANCE

The behavior of circuits and circuit elements is described in terms of the relationship between voltages and current at their terminals. This is most commonly done by specifying resistance, reactance, and impedance at the terminals. As we have seen, these three parameters are ratios of voltage to current and are measured in ohms. When circuit elements are connected in series, the total circuit impedance can be found by a vector addition of resistances and reactances. When elements are connected in parallel, finding the total circuit impedance is more complicated.

With the parallel connection it would be much easier if we were to use the reciprocals of resistance, reactance, and impedance. These are *conductance*, *susceptance*, and *admittance*, respectively, and they are measured in *mhos*. These reciprocal quantities are ratios of current to voltage. Just as we could say that one ohm equals one volt per ampere, we could say that one mho equals one ampere per volt.

Disadvantages of Admittance

The reason admittance is not used more widely in broadcast work is twofold. In the first place, component values are traditionally specified in ohms. Although it might be easier to use mhos to solve a problem, by the time we have converted everything to mhos we have done as much work as if we had solved the problem using ohms. For example, suppose we want to find the total resistance of a 2-ohm and an 8-ohm resistor connected in parallel. The conventional approach is to take the product of the two and divide it by their sum.

$$R = \frac{2 \times 8}{2 + 8} = \frac{16}{10} = 1.6 \text{ ohms}$$

We can find the total conductance of two conductances in parallel by simply adding them together. To do this in the above case, we convert each resistance to a conductance, then add them together

$$G = 1/2 + 1/8 = 0.5 + 0.125 = 0.625 \text{ mho}$$

Then, to use the result in most applications, we would have to convert the conductance back to a resistance. If we had to make the computation by hand, we would gain nothing. Fortunately, with an electronic calculator the computation is very simple. Thus we can use the concept of admittance whenever it will either simplify computation or make things clearer.

The only remaining obstacle to using admittance, conductance, and susceptance is that the magnitudes, being unfamiliar, are apt to be meaningless. For example, most broadcast engineers wouldn't realize immediately that an admittance of 20 millimhos was the same as an impedance of 50 ohms.

Inasmuch as the concepts are very useful, we will take a few minutes to review their meaning and the techniques for using them.

Conductance and Susceptance

Conductance is the reciprocal of resistance and is usually represented by the symbol G. That is,

$$G = \frac{1}{R}$$

Thus, a resistance of 5 ohms would correspond to a conductance of 0.2 mho or 200 millimhos.

Susceptance is the reciprocal of reactance and is usually represented by the symbol B. The formulas for inductive and capacitive susceptance are

$$B_C = 2\pi fC \qquad B_L = \frac{1}{2\pi FL}$$

Because susceptance is the reciprocal of reactance, *inductive susceptance* has a minus sign, whereas *capacitive reactance* has a minus sign. Thus the total susceptance in a parallel circuit is $B_C - B_L$.

Admittance Y is a complex number that represents the ratio of current to voltage. It includes both conductance and susceptance, and in a parallel circuit it is equal to

$$\mathbf{Y} = G + jB$$

The method of converting resistance to conductance, or reactance to susceptance, is straightforward. We simply take the reciprocal of one to get the other. When it comes to converting between impedance and admittance, there is often a great deal of confusion. For example, suppose that we have a resistance of 4-ohms in series with an inductive reactance of 3 ohms. We know that the magnitude of the impedance is 5 ohms. The impedance may be expressed as

$$\mathbf{Z} = R + jX = 4 + j3 = 5 \text{ ohms}$$

Now suppose that we want to find the admittance looking into this circuit. Since $1/4 = 0.25$ and $1/3 = 0.33$, there is a temptation to say that the admittance looking into the circuit is

$$\mathbf{Y} = G + jB = 0.2 + j0.33 \text{ mho}$$

This temptation should be resisted, because the expression is wrong. To find the right way to make the conversion, we must take a look at what our equations mean.

When we write the expression $\mathbf{Z} = R + jX$, we are stating that the impedance between two wires, such as the leads of the box in Fig. 1-27A, is the same as the impedance of a circuit consisting of a resistance R connected *in series with* a reactance X. When we write the expression $\mathbf{Y} = G + jB$, we are saying that the admittance between two wires, such as the leads of the box in Fig. 1-27B, is the same as the admittance seen across a circuit containing a conductance G connected *in parallel with* a susceptance B. The two circuits are not the same. If we want to find the total admittance of two elements in series, we must take their sum over their product, just as with resistors in parallel.

When we are concerned only with admittance and impedance, and when they are expressed in polar form, we

Fig. 1-27. Conductances and susceptances connected in series (A) and parallel (B).

can simply take the reciprocal of one to get the other, that is, divide the vector quantity into 1. Thus if we want to find the admittance corresponding to an impedance of $2 \angle 30^\circ$ ohms, we calculate

$$G = \frac{1}{Z} = \frac{1}{2 \angle 30^\circ} = 0.5 \angle -30^\circ \text{ mho}$$

Chapter 2
Principles of Transmission Lines

It is necessary to locate an antenna at some distance from the transmitter. In television and FM stations the antennas are located on tall towers to get good coverage; in AM directional stations, antennas consist of two or more widely separated towers. It is necessary to get the signal to the antenna with a minimum of loss and with as little radiation as possible along the way. Transmission lines of one type or another are used for this purpose. In this chapter we consider properties of transmission lines that are fundamental and apply equally to all types of broadcast antennas. Later we will consider feeder systems for particular antenna types.

One of the most important requirements for a transmission line is that it must not radiate signals. Radiation patterns can be controlled best at the antenna itself. If the line should radiate, it would not only waste energy, but it might radiate energy in such a direction as to defeat the directional design of the antenna. A directional antenna is designed to radiate a minimum amount of energy in certain directions, to "protect" areas served by other stations on the same frequency. If the transmission line radiated, it might put an interfering signal in the protected area.

Whenever an RF current flows in a wire more than about 1/10 wavelength long, the wire will tend to act as an antenna and radiate energy. In a transmission line this tendency toward radiation is minimized by using closely spaced conductors in which currents are flowing in opposite directions (Fig. 2-1). Inasmuch as the currents in the two conductors are in opposite directions, the fields from them will also be in opposition; and at a distance from the line, the fields will tend to cancel.

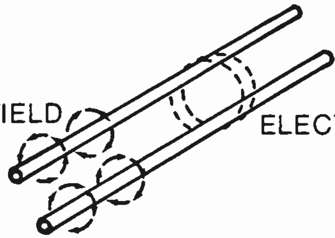


Fig. 2-1. Fields around a transmission line.

Field cancellation is fundamental to the operation of transmission lines, antennas, and all forms of shielding. We can assume that whenever an electric charge moves, it will tend to make every other charge in the universe move at the same frequency. If other charges in the same area move in the opposite direction, they will have equal and opposite effects on still other charges in the universe. This is how a coaxial cable minimizes radiation. The field from the outer conductor cancels the field from the inner conductor. It can be shown mathematically that this cancellation takes place right at the outer conductor. The outer conductor does *not* confine the field in the way a water pipe contains water.

IDEAL LINE

It is easier to gain insight into the operation of many practical devices by first considering an ideal model, studying its behavior, and then modifying it so that it more closely resembles a practical device. We do this with transmission lines by starting out with an idealized line. We assume that it consists of two parallel conductors that have no series resistance and no leakage between them.

Having decided to neglect resistance, we can almost intuitively draw the equivalent circuit. Since each of the wires has appreciable length, it has inductance. We know, therefore, that there will be series inductances in our circuit. Since the two conductors are in close proximity, we know that there will be capacitance between them. It is not surprising, therefore, to find that the ideal transmission line has an equivalent circuit like that shown in Fig. 2-2A. We can, without serious error, further simplify the circuit by placing all of the inductance in one conductor as in Fig. 2-2B.

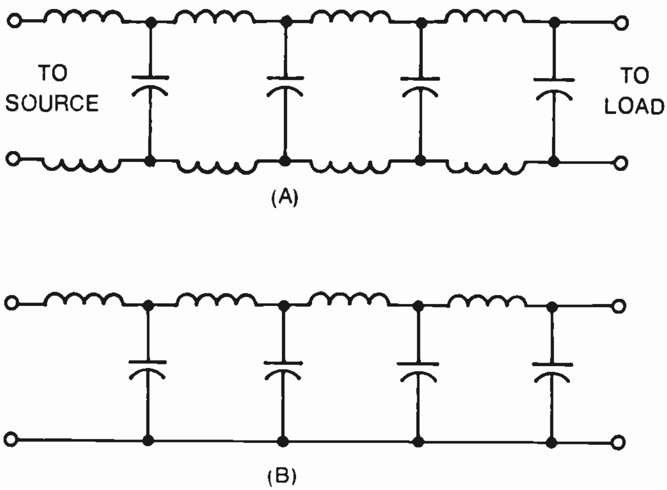


Fig. 2-2. Equivalent circuit of an ideal transmission line.

Of course, in an actual line the inductance and capacitance are distributed uniformly along the line, and not lumped as shown. Nevertheless, the equivalent circuit very closely approximates an actual line. In dealing with this circuit, we will not consider each of the inductances and capacitances separately, but will deal with the inductance and capacitance per unit length of the line. That is, we will use units like henries per foot and farads per foot.

CHARACTERISTIC IMPEDANCE

One very important property of any transmission line is its *characteristic impedance*. We can best understand this term by considering a fictitious line that is infinitely long. Let us connect a battery to this ideal line through a switch, as shown in Fig. 2-3. Current flows when the switch is closed. All of the capacitances along the line have no charge in them before the switch is closed, so current rushes in to charge them. The current is not infinitely large, however, because it is limited by the series inductances.

As a matter of fact, there is a definite relationship between the applied voltage and the resultant current that depends only on the construction of the line itself. Since the line is assumed to be infinitely long, a steady current flows. No matter how many fictitious capacitors become charged, there are always more to charge. What this means is that energy is flowing into the line, where it is stored in electric and magnetic fields. Since the energy is continuously flowing from left to right in the figure, and not returning, the voltage and current are in phase, and the line "looks like" a resistance.

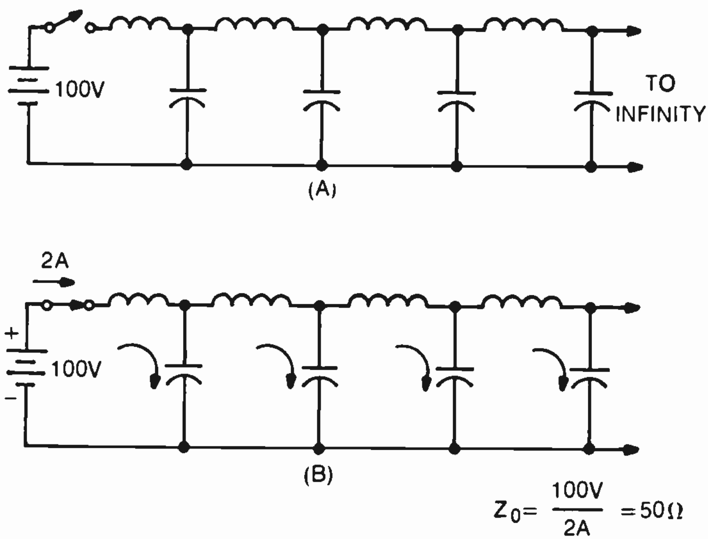


Fig. 2-3. Impedance of an ideal transmission line.

In circuit theory we call the ratio of voltage to current *impedance*, and since there is a definite relationship between the applied voltage and the resulting current in our infinitely long line, we can say that it has a *characteristic impedance*. In our example in Fig. 2-3 the applied voltage is 100V, and the resulting current is 2A; we can say that the characteristic impedance Z_0 of the line is

$$Z_0 = \frac{100V}{2A} = 50 \text{ ohms}$$

The voltage and current in the line are in phase. This means that the characteristic impedance of our ideal line is a pure resistance of 50 ohms.

As long as the ideal transmission line is infinitely long, it looks just like a resistor. There would be no way of telling by electrical measurements whether the battery in Fig. 2-3 is connected to an infinite line having a characteristic impedance of 50 ohms or to a 50-ohm resistor.

In Fig. 2-4 we have the same 50-ohm line connected to a 100V battery. Suppose we were to cut the transmission line at the line A-A. Inasmuch as the transmission line is said to be infinite, the remaining infinite section to the right of the cut must still look electrically like a 50-ohm resistor. We can therefore cut the line and terminate it in a 50-ohm resistance, as shown in Fig. 2-4B, and it will still look like a 50-ohm resistor at the input terminals. The line is then said to be *terminated in its characteristic impedance*. The input impedance of a line so terminated equals its characteristic impedance. This is true

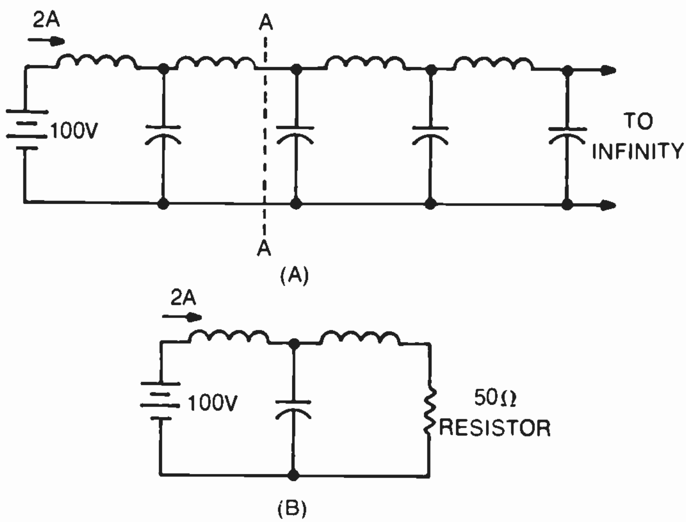


Fig. 2-4. Termination of a transmission line.

regardless of the length of the line, as long as we neglect any losses. When a line is terminated in some value other than its characteristic impedance, its input impedance will depend on the value of the terminating impedance and the length of the line, as well as on its characteristic impedance.

In some older literature the characteristic impedance of a transmission line is called the *surge* impedance, because it is the ratio between the applied voltage and the current that would surge into the line if it were infinitely long.

The value of the characteristic impedance of a transmission line depends entirely on its physical construction. In a lossless line the characteristic impedance is given by

$$Z_0 = \sqrt{\frac{L}{C}}$$

where L is the inductance in henries per unit length, and C is the capacitance in farads per unit length. Any units of length may be used as long as they are the same in both cases.

The characteristic impedance can also be expressed in terms of the physical dimensions of the line. In a 2-wire open line

$$Z_0 = 276 \log \frac{d}{2D}$$

where d is the diameter of the conductors, D is their spacing, and both are expressed in the same unit.

In a coaxial cable in which the space between conductors is filled with air, the characteristic impedance is given by

$$Z_0 = 138 \log \frac{D}{d}$$

where d is the diameter of the inner conductor, D is the diameter of the outer conductor, and both are expressed in the same unit.

A plot of characteristic impedance as a function of line dimensions is given in Fig. 2-5.

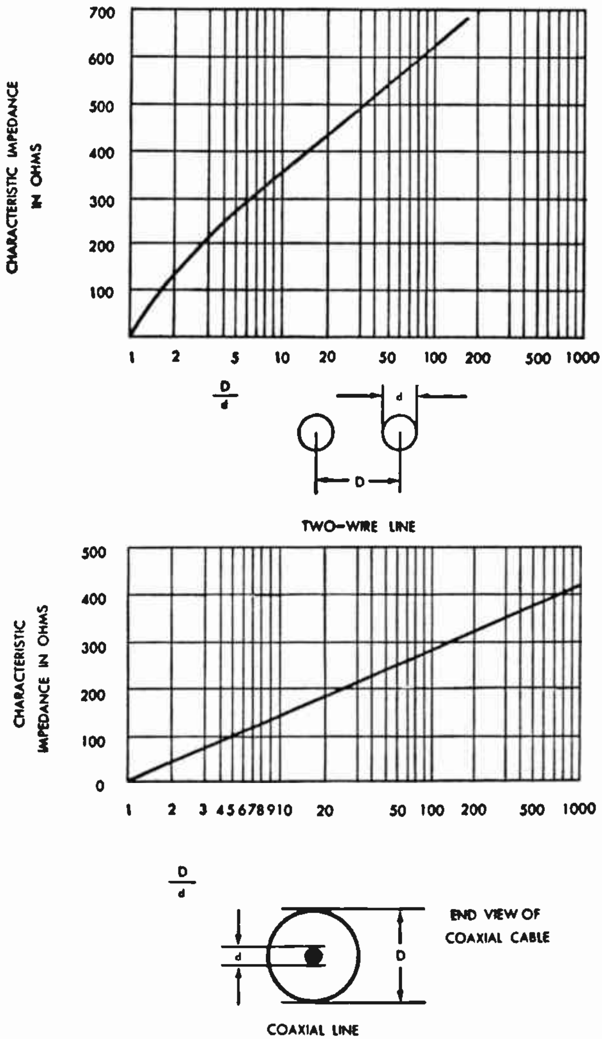


Fig. 2-5. Characteristic impedance graphs for 2-wire and coaxial ines.

REFLECTIONS

We are not always fortunate enough to have all of the transmission lines we work with terminated in their characteristic impedance. In standard broadcast work this is our goal, but in FM and TV we actually use the properties of transmission lines that are not terminated in this way to produce changes in impedance level.

The two extreme cases of lines that are not terminated in their characteristic impedance is when the far or receiving end of the line is either open or shorted. By studying these two cases, we can gain some insight into how lines behave with other values of terminating impedance. Let us first consider the case when the receiving end is open.

Open Transmission Line

Figure 2-6A shows a source connected through a switch to a line. Both the internal impedance of the source and the characteristic impedance of the line are pure 50-ohm resistances. For simplicity let us say that our source produces a DC voltage and that its open-circuit voltage is 100V. When the switch is closed, current rushes into the line to charge its distributed capacitance. Before the energy reaches the end of the line, the source has no way of "knowing" that the receiving end is open, so the line behaves just like a 50-ohm resistor. Thus, during this time, the voltage from the source divides evenly between its own internal impedance and the impedance of the line. There is 50V across the line and 50V across the internal impedance of the source. Thus a 50V wave will propagate along the line toward the receiving end. Bear in mind that the only way the energy gets to the end of the line is by being stored in the electric and magnetic fields associated with the conductors.

Just as the voltage wave reaches the end of the line, there is a current of 1A flowing in the equivalent inductance L of the last section of the line. This current charges the last capacitor C to a voltage of 50V. When this happens, the current stops abruptly; there is no place for it to go. Current is needed to sustain the magnetic field associated with L, and when the current drops to zero, the field collapses. The collapsing field, in turn, induces a voltage of such polarity as to increase the voltage across C, as shown in Fig. 2-6B. All of the energy that was stored in the magnetic field of L is transferred to C, and this is just enough to double the voltage to 100V. Thus a 50V wave propagates back toward the source, as shown in Fig. 2-6C, raising the voltage across the capacitor to 100V. The current involved is 1A because of the characteristic impedance of the line. When the first capacitor in the line charges to 100V, the whole line is charged to 100V. There is now no current flowing in the line and no voltage drop across the internal impedance of the source, as shown in Fig. 2-6D. Under this condition all of the energy in the line is stored in its capacitance.

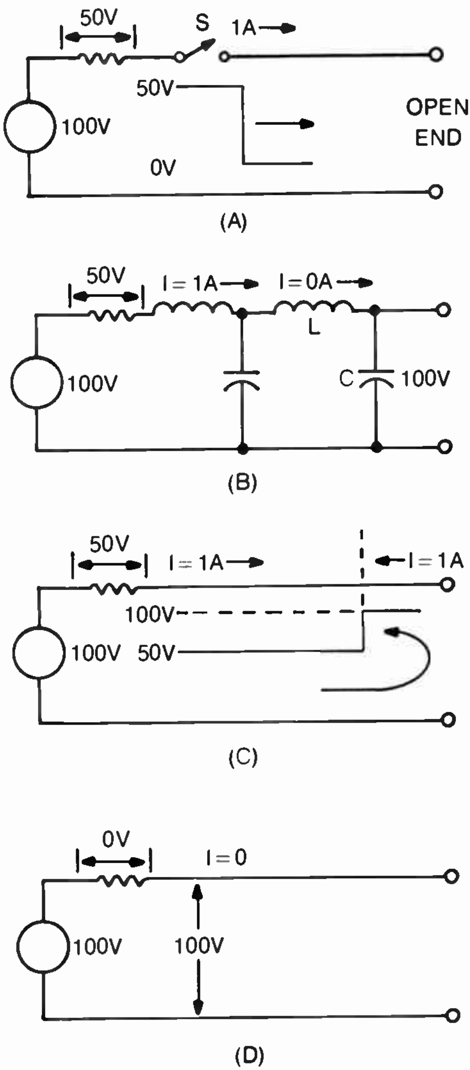


Fig. 2-6. Reflection from an open transmission line.

There are two points worth noting about what happens when a signal reaches the *open* end of a transmission line:

1. The voltage wave is reflected, in phase, with no change in waveform.
2. The current is reflected, in the opposite phase, with no change in waveform.

Shorted Transmission Line

The situation where the receiving end of the line is shorted is shown in Fig. 2-7. In Fig. 2-7B, because the end of the line is shorted, we end the equivalent circuit with an inductance (a capacitor connected across the shorted end of the line would have no effect). As in the case where the end of the line was open, when the switch is closed a voltage wave of 50V travels down the line. Just ahead of the wave the voltage across the line is zero, and just behind the wave it is 50V. At the instant the wave reaches the end of the line (Fig. 2-7B), there is nothing to limit the current in the inductor, and it increases until the induced voltage is just high enough to reduce the voltage across the last capacitor to zero. Note that the current is in the same direction, but that the induced voltage has the opposite polarity of that traveling down the line from the source. A very short time later the voltage rises across the preceding inductance and discharges the next capacitor. Thus there is a reflected wave of 50V that is out of phase with the original wave, as shown in Fig. 2-7C. This reflected wave reduces the voltage across the line to zero. The reflected wave of current is equal in magnitude to the original current, and since it is in phase with the original current, the current in the line doubles. After the reflected wave reaches the source, the current in the line is 2A; there is no voltage across the line, and all of the voltage drop is across the internal impedance of the source (Fig. 2-7D). The energy in the line is then all stored in its inductance.

There are two points worth remembering about what happens when a signal reaches the *shorted* receiving end of a transmission line:

1. The voltage wave is reflected, out of phase, with no change in waveform.

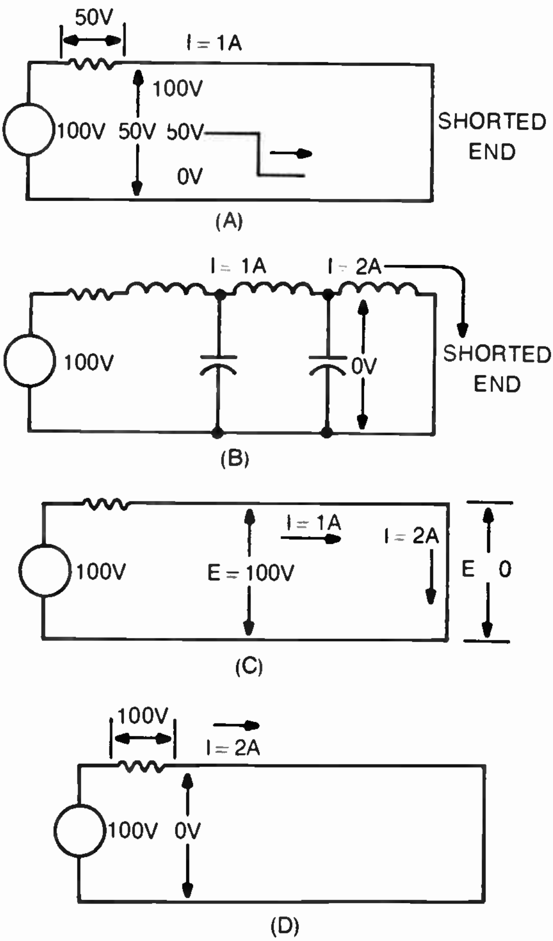


Fig. 2-7. Reflection from a shorted transmission line.

2. The current is reflected, in phase, with no change in waveform.

RF Signals on Transmission Lines

We use transmission lines to carry RF signals, not DC voltages as in the preceding examples. The reflection action is exactly the same for either RF or DC signals at any instant, but in antenna work we are not interested in instantaneous phenomena. Rather we are interested in the steady-state behavior of transmission lines. Because RF signals are periodic sinusoidal voltages and currents, the manifestation of reflection will be considerably different than in the DC case we have just looked at.

STANDING WAVES

When an RF signal reaches the open end of a transmission line, the voltage is reflected, in phase with the incident voltage (Fig. 2-8). The *reflected* wave is what we would get if we folded the forward or *incident* wave back on itself. The actual voltage distribution along the line is the sum of the incident and reflected waves. Although the incident wave is moving to the right in the figure, and the reflected wave is moving to the left, the sum of the two will be a wave that doesn't move at all along the line. It is called a stationary or *standing* wave.

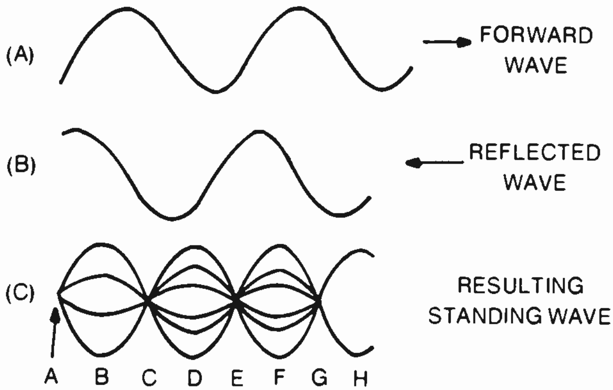


Fig. 2-8. Standing waves.

At points *B*, *D*, *F*, and *H*, of Fig. 2-8, the voltage varies between a maximum positive and maximum negative value. At points *A*, *C*, *E*, and *G*, the incident and reflected waves cancel completely at all times, so the voltage at these points is zero. In a practical line the voltages do not actually reach zero, but some other minimum value.

The points of maximum voltage of a standing wave are usually called voltage *loops*, and the points of minimum voltage are called voltage *minima*, or *nulls*.

It is easier to get a good feeling for how standing waves are formed by considering the behavior of a rope that is tied securely at one end (Fig. 2-9). When the rope is given a shake, a wave travels along it toward the far end, which is secured. When the wave reaches the end of the rope, it is reflected back along the rope, and it gives a jerk when it gets back to the shaker's hand. If the rope is shaken rapidly, waves travel forward and backward along the rope at the same time. If the shaking is done at the proper rate, a standing wave is formed on the rope, as shown in the figure. This situation is analogous to the formation of standing waves on a transmission line. In both cases, a standing or stationary wave is formed by the sum of two waves of the same frequency moving in opposite directions.

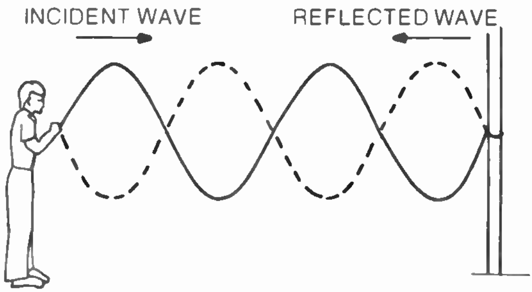


Fig. 2-9. Standing wave on a rope.

If we were to measure the voltage of a standing wave along a transmission line with an RF voltmeter and plot the indications of the meter as a function of distance along the line, we would get a plot like that of Fig. 2-10. The indication of the meter would be proportional to the rms value of the voltage at each point and would not show the instantaneous value or polarity of the voltage. Standing waves are usually plotted in this way.

Standing Waves for Various Terminations

Inasmuch as a standing wave on a transmission line is caused by a reflection, which in turn is caused by a mismatch at the receiving end, the nature of the standing wave depends on the way in which the line is terminated. Figure 2-11 shows several different terminations and the resulting standing-wave patterns.

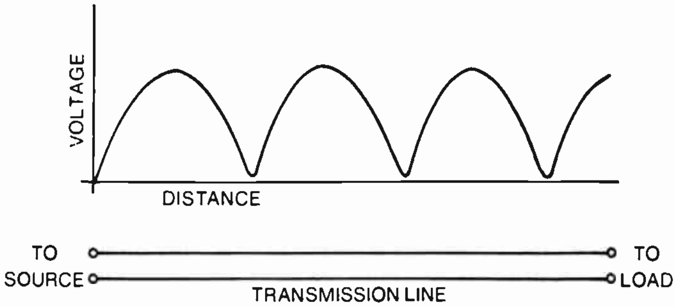


Fig. 2-10. Plot of a standing wave along a transmission line.

In Fig. 2-11A the line is terminated in its characteristic impedance, and since there is no reflection in this case, there is no standing wave. In Fig. 2-11B we have the same situation, but there is some loss in the line. Although there is no reflection and no standing wave, the voltage drops along the line because of the losses in the line.

Figure 2-11C shows the standing-wave pattern that results when the receiving end of a transmission line is open. At the receiving end of the line, the voltage is maximum and the current is minimum. This is what we might expect, since with an open circuit at the end of a line, there is no place for the current to flow. The standing-wave pattern resulting when the receiving end of a line is shorted is just the opposite (Fig. 2-11D). The voltage at the receiving end of the line is now minimum and the current is maximum. This, again, is what we might expect because there can be no voltage across a short circuit. In the cases of open and shorted lines, the peak value of the standing wave will be twice the peak value of the incident voltage.

Whenever a transmission line is terminated in anything except a resistance equal to the characteristic impedance of the line, there will be a reflection and, consequently, a standing wave on the line. If the termination is a resistance that is higher or lower than the characteristic impedance of the line, some of the energy in the incident wave will be absorbed in the termination; but since the resistance is not equal to the characteristic impedance of the line, some of the energy will also be reflected.

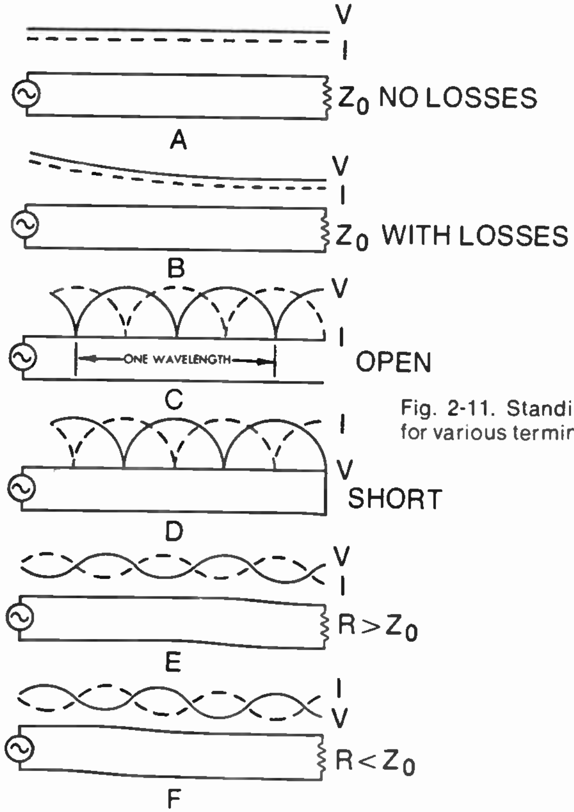


Fig. 2-11. Standing-wave pattern for various terminations.

The voltage of the reflected wave can be found from the formula

$$V_R = V_F \frac{R_L - Z_0}{R_L + Z_0}$$

where V_R = reflected voltage
V_F = forward voltage
R_L = terminating resistance
Z_0 = characteristic impedance

Inspection of this equation shows that the reflected voltage V_R will always be equal to or less than the forward voltage V_F. It cannot be greater than the forward voltage.

Voltage Standing-Wave Ratio

A common measure of the magnitude of a standing wave on a transmission line is the *voltage standing wave ratio*, or

VSWR. It is the ratio of the maximum voltage on the line to the minimum voltage; that is,

$$VSWR = \frac{V_{max}}{V_{min}}$$

The voltage standing wave ratio is also numerically equal to the ratio of the terminating resistance to the characteristic impedance of the line, or vice versa. It is usually arranged so that it will be a number greater than 1. Thus

$$VSWR = \frac{R_L}{Z_0} \quad \text{or} \quad \frac{Z_0}{R_L}$$

The voltage standing wave ratio is based on the maximum and minimum values of voltage on a transmission line. We might just as well have specified a *current* standing wave ratio; it would have the same numerical value. In fact, much of the time, the ratio is specified simply as *standing wave ratio*. The reason for using the voltage standing wave ratio is that it is usually easier to measure the voltage on a line than the current.

In general, a value of terminating resistance that is small compared with the characteristic impedance of the line will cause a standing wave pattern that is similar to that from a short circuit, except that the standing wave isn't as large. Similarly, with a resistance that is higher than the characteristic impedance of the line, the standing wave pattern will be similar to that resulting in an open circuit. Again, the standing wave will not be as large.

In Fig. 2-11, E and F show the standing-wave patterns that result from resistive terminations that are higher and lower than the characteristic impedance of the line. The size, or magnitude, of the standing wave is a measure of how much the termination deviates from the characteristic impedance. The closer the value of the terminating resistance to the characteristic impedance, the smaller the standing wave.

Inductive and Capacitive Terminations

So far we have considered the termination of a transmission line to be a pure resistance. We found that if the termination were equal to the characteristic impedance of the line, all of the energy would be absorbed by the load. If the termination had any other value, some of the energy would be reflected back toward the source. Let's look at what would happen if the termination were a pure inductance or capacitance. We know that reactive elements such as inductances and capacitances only store energy and do not dissipate it; therefore, in a lossless line with a reactive termination, all of the energy will be reflected back toward the source.

The exact nature of a reflection from a reactive termination depends on the value of the reactance. In Fig. 2-12 the terminations have a reactance that is numerically equal to the characteristic impedance of the line.

Figure 2-12A shows a capacitive termination. The capacitive reactance and the resistive characteristic impedance of the line form a 45° phase-shifting network. The reflected voltage is shifted 45° in one direction, and the reflected current is shifted 45° degrees in the opposite direction. The standing-wave pattern is such that the voltage is maximum when the current is minimum, and voltage peaks are separated from current peaks by 90°, or 1/4 wavelength along the line.

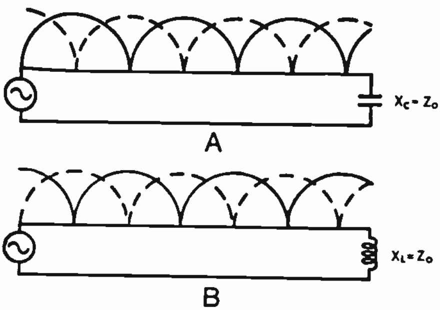


Fig. 2-12. Standing-wave pattern for reactive terminations.

Figure 2-12B shows an inductive termination. The pattern is the same as in the capacitive case except at the load. Here the phases of voltage and current are just the opposite. This is because the phase shift of the voltage in an inductance is just the opposite of that in a capacitance.

DRIVING-POINT IMPEDANCE FOR VARIOUS LINES

We have found that every transmission line has a characteristic impedance that depends only on the dimensions of the line, and not on its length or what may be connected to it. When the line is terminated in its characteristic impedance, the *driving-point impedance*—that is, the impedance seen looking into the sending end—is equal to the characteristic impedance of the line. We have also seen that with any other termination, of any type, there are standing waves on the line. Since impedance is the ratio of voltage to current, we can see that whenever there is a standing wave, the impedance varies along the line.

Figure 2-13 shows an arbitrary length of transmission line, terminated in a short circuit, together with a plot of standing waves of voltage and current that exist on the line. At the termination the voltage is zero because a voltage cannot exist across a short circuit. The current at this point is maximum because current is maximum through a short circuit. The voltage reaches a maximum value at a point 1/4 wavelength from the load and then drops back to zero at a point 1/2 wavelength away from the load. Thus, if we cut the line at a point 1/4 wavelength back from the shorted termination (*A-A* in Fig. 2-13), the driving-point impedance would be very high. The line would for all practical purposes look like an open circuit. If the line was cut 1/2 wavelength from the load, it would have a very low impedance and would look like a short circuit. At other fractions of a wavelength the line would look like something between a short and an open circuit. We are now considering only lines with no losses, and a short-circuit termination, which doesn't absorb any energy. Now we can draw another conclusion about the driving-point impedance of such a line: the driving-point impedance seen looking into such

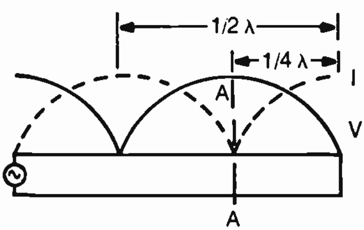


Fig. 2-13. Standing waves on shorted transmission line.

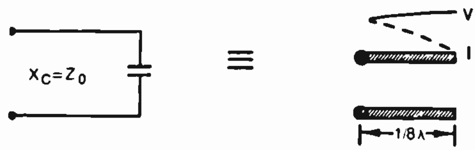
a line is either zero, infinity, or some value of *reactance*; that is, the voltage and current at the sending end will be 90° out of phase. A resistive value of driving-point impedance would mean that energy was being dissipated somewhere rather than simply being stored in the line.

In studying the behavior of open and shorted lines, it is useful to consider an element of a line that is 1/8 wavelength. Figure 2-14A shows a 1/8-wavelength line that is open at the receiving end. Since the end is open and the line is not very long, very little current will flow in it, and therefore practically no magnetic field will exist around the wires. This is equivalent to saying that there is little or no inductance. On the other hand, there will be a substantial voltage between the wires, so a substantial amount of energy may be stored in the electric field. This is the same as saying that the line has capacitance. Thus an open-ended 1/8-wavelength section of transmission line looks electrically like a capacitor at its driving point. That is, the driving-point impedance is capacitive.

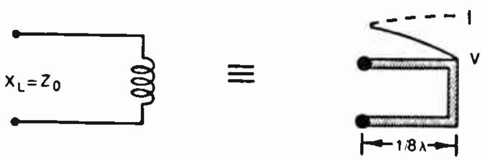
We can intuitively get a good idea of the magnitude of this capacitive reactance. We know that the voltage and current are shifted 90° in a 1/4-wavelength line, so it is logical to suspect that they will be shifted 45° (in opposite directions) in a 1/8-wavelength line. Thus the 1/8-wavelength open-ended line will behave exactly the same as a capacitive termination that produces a 45° phase shift.

From circuit theory we know that a 45° phase shift is produced in an *RC* circuit where the capacitive reactance is numerically equal to the series resistance. In a 1/8-wavelength line that is open at the receiving end, the capacitive reactance seen looking into its terminals is numerically equal to the characteristic impedance of the line. Thus, an open-ended 1/8-wavelength section of a 50-ohm line will have a driving-point impedance equal to a capacitive reactance of 50 ohms.

Now, let's look at a 1/8-wavelength section of line in which the receiving end is shorted (Fig. 2-14B). Here, since the end is shorted and the line isn't very long, there is not much voltage drop across it or, consequently, much energy stored in the electric field. This means that it will have little or no capacitance. On the other hand, since the far end is shorted, there is a large current, and quite a bit of energy is stored in the magnetic field. This means that inductance will predominate. By the same reasoning that we used in connection with Fig. 2-14A, we can conclude that a shorted 1/8-wavelength section of line will look like an inductive reactance equal to the characteristic impedance of the line. Thus a 1/8-wavelength section of 50-ohm line shorted at the receiving end will have a driving-point impedance equal to an inductive reactance of 50 ohms.



(A) OPEN-ENDED 1/8-WAVELENGTH SECTION



(B) SHORTED 1 8-WAVELENGTH SECTION.

Fig. 2-14. Conditions in a 1/8-wavelength line.

QUARTER-WAVE SECTION

One of the most interesting lengths of transmission line is the quarter-wave (1/4-wavelength) section. This line inverts the impedance in which it is terminated. The shorted quarter-wave line may be thought of as an open 1/8-wavelength section feeding a shorted 1/8-wavelength section, as shown in Fig. 2-15A. This is equivalent to putting a capacitance and an inductance in parallel, so the shorted quarter-wave line looks like a parallel-resonant circuit (Fig. 2-15B)—that is, it has a very high driving-point impedance. If there was no loss in the line, the driving-point impedance would be infinite. In practical lines with some loss, the driving-point impedance of a shorted quarter-wavelength section is not infinite, but very high.

The shorted quarter-wavelength section resembles a parallel-resonant circuit in other ways. The driving-point impedance is a capacitive reactance at frequencies above resonance, and an inductive reactance at lower frequencies.

When the quarter-wave line is open at the receiving end, the voltage at the end is high and the current low, as we would expect with an open circuit. At the sending end—which is quarter wavelength, or 90° away—both voltage and current have changed by 90°. This means that at the sending end the current is high and the voltage is zero, so at the sending end the line looks like a short circuit. Of course, since all of the energy is reflected from the open end, there is no dissipation of energy. The only type of circuit that stores energy in both the electric and magnetic fields, and yet looks like a short circuit at one frequency, is the series-resonant circuit. The open quarter-wave line does, indeed, look electrically like a series-resonant circuit, as shown in Fig. 2-15C.

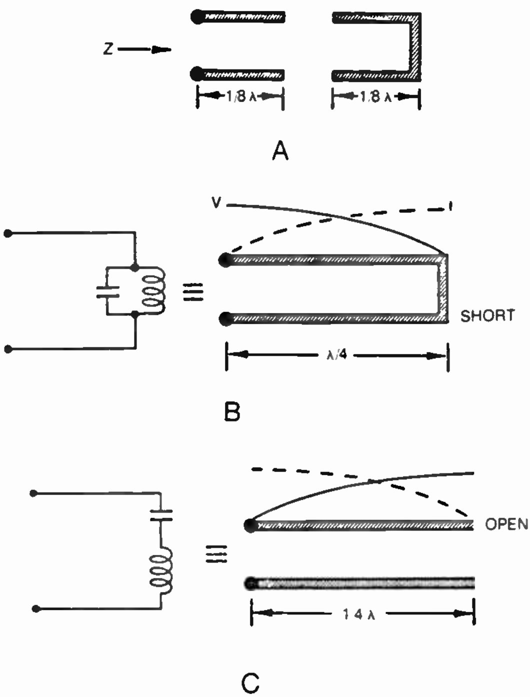


Fig. 2-15. Quarter-wave section of transmission line.

Thus, from the sending end, a shorted quarter-wave line looks like an open circuit, and an open quarter-wave line looks like a short circuit. This impedance inversion takes place with any value of termination except a resistance equal to the characteristic impedance of the line. When a quarter-wave line is terminated in a resistance greater than its characteristic impedance, its driving-point impedance is a resistance that is smaller than its characteristic impedance, and vice versa. The mathematical relationship between the characteristic impedance Z_0 of the quarter-wave line, its terminating impedance Z_L, and its driving-point impedance Z_{in} is given by

$$\frac{Z_{in}}{Z_0} = \frac{Z_0}{Z_L}$$

which can also be written

$$Z_{in} = \frac{Z_0^{\ 2}}{Z_L}$$

Thus a quarter-wave section of 50-ohm line terminated in a 25-ohm resistor will have a driving-point impedance of

$$Z_{in} = \frac{50^2}{25} = \frac{2500}{25} = 100 \text{ ohms}$$

Using an equation given earlier, we can calculate the reflected voltage to be one-third of the incident voltage and opposite in sign. With this information we can plot the standing wave on the line (Fig. 2-16).

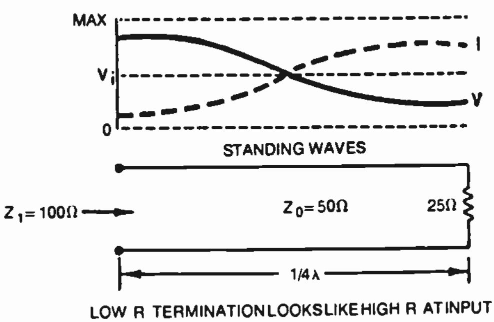
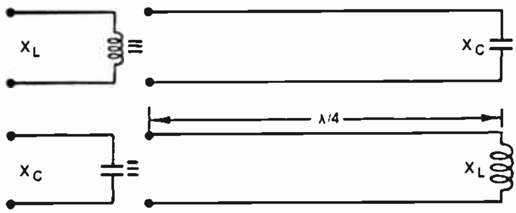


Fig. 2-16. 1/4λ section of 50Ω terminated in 25Ω.

The quarter-wave line can be used as an impedance-matching transformer *at a single frequency.* This is of little interest in standard AM broadcasting, but it can be used to advantage in FM and TV antenna systems.

The equation given for the driving-point impedance of a quarter-wave line can be used to show that a quarter-wave line also inverts reactance. A quarter-wave line that is terminated in a capacitance has an inductive reactance at its driving point, and vice versa (Fig. 2-17).



CAPACITANCE LOOKS LIKE INDUCTANCE AND VICE VERSA

Fig. 2-17. Reactance inversion.

THREE-EIGHTHS-WAVELENGTH LINE

The impedance of the three-eighths-wave line can be found as easily as that of the quarter-wave line. It consists of two sections that we are already familiar with—the one-eighth-wave line and the quarter-wave line. Consider first the three-eighth-wave line that is shorted at the receiving end (Fig. 2-18 A and B). The shorted eighth-wave section will look like an inductive reactance equal numerically to the characteristic impedance of the line. This inductive reactance then terminates the quarter-wave section, which will invert the impedance so it will look like a capacitive reactance—again, numerically equal to the characteristic impedance of the line.

When the three-eighth-wave line is open at the receiving end, as shown in Fig. 2-18C, the driving-point impedance is an inductive reactance equal to the characteristic impedance of the line.

HALF-WAVE LINE

The half-wave line is of interest because it is used in FM and TV antenna feeder systems and because it can lead to

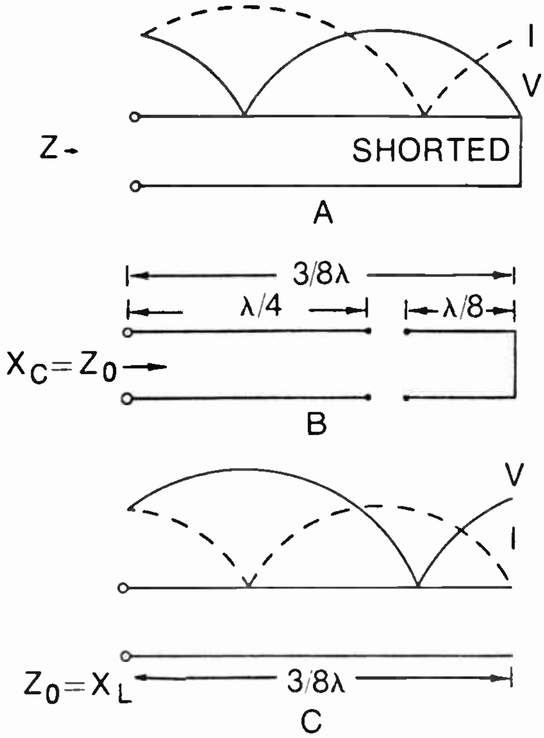


Fig. 2-18. Three-eighths-wavelength section.

confusion when a line in an AM feeder system happens to be approximately half-wavelength long. The operation of the half-wave line is easy to see, because it consists of two quarter-wave sections connected together, as shown in Fig. 2-19. The quarter-wave section nearest the termination inverts its impedance. The next quarter-wave section inverts the impedance again bringing it back to its original value. Thus the driving-point impedance of the half-wave line is exactly equal to the terminating impedance.

SUMMARY OF TRANSMISSION-LINE IMPEDANCES

We use the term *impedance* in three separate senses when working with transmission lines:

1. The *characteristic* impedance Z_0 of the line itself depends only on the physical construction of the line,

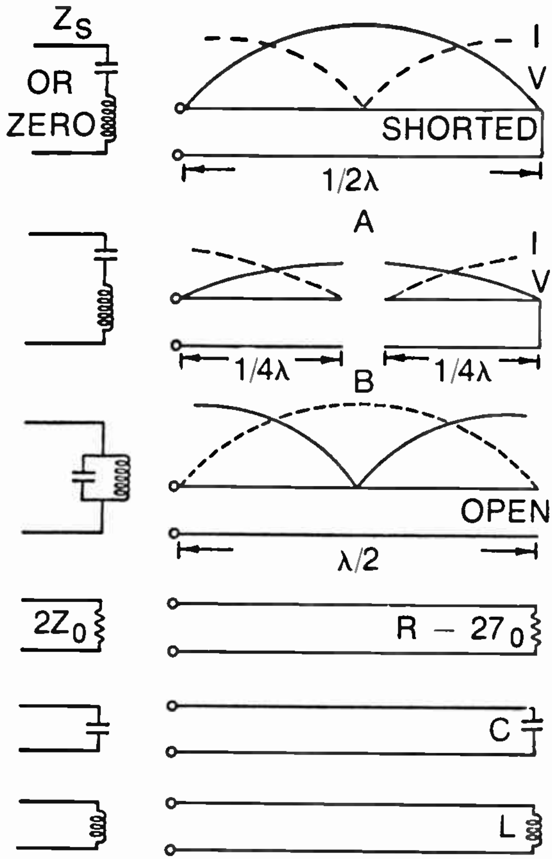


Fig. 2-19. Half-wave section.

and not on its length or on what might be connected to it.

2. The *terminating or load* impedance, Z_L is the impedance that is connected to the receiving end of the line.
3. The *driving-point* or *sending-end* impedance Z_{in} is the impedance seen looking into the sending end of the line. When the load impedance is equal to this characteristic impedance, the driving-point impedance is equal to the characteristic impedance. With any other value of load impedance, the driving-point impedance will depend on the load impedance, the characteristic impedance, and the length of the line.

We can get a rough idea of the driving-point impedance for almost any line length and termination from the situations we looked at on the preceding pages. These can be summarized as follows:

1. Open or shorted eighth-wave sections of line have a driving-point impedance that is reactive and numerically equal to the characteristic impedance of the line.
2. Quarter-wave sections invert the impedance connected to the receiving end.
3. Quarter-wave sections act like resonant circuits.
4. Half-wave sections have a driving-point impedance equal to the terminating impedance.

Figure 2-20 shows whether the impedance along an open or shorted line is inductive, capacitive, or resistive. These charts are based on low- and high-resistance terminations, rather than on short and open terminations, because, in practice, we can get neither a perfect short circuit nor a perfect open circuit.

VELOCITY OF PROPAGATION

In the preceding sections of this chapter, we have considered the lengths of transmission lines in fractions of a wavelength. We have ignored the velocity of propagation of a signal in a particular line, on which depends the length of a wave.

The wavelength of a signal—that is the physical length of one wave of the signal in space—is given by

$$\lambda = \frac{c}{f}$$

where c is the velocity of propagation—300,000,000 meters per second—and f is the frequency in hertz. If a signal were to travel at this same velocity in a transmission line, we could use this formula to find the physical length of a wavelength of transmission line. But, more often, the velocity is somewhat lower in a transmission line than in free space.

There is a relationship between the characteristics of free space and the velocity of propagation in it that will give us a

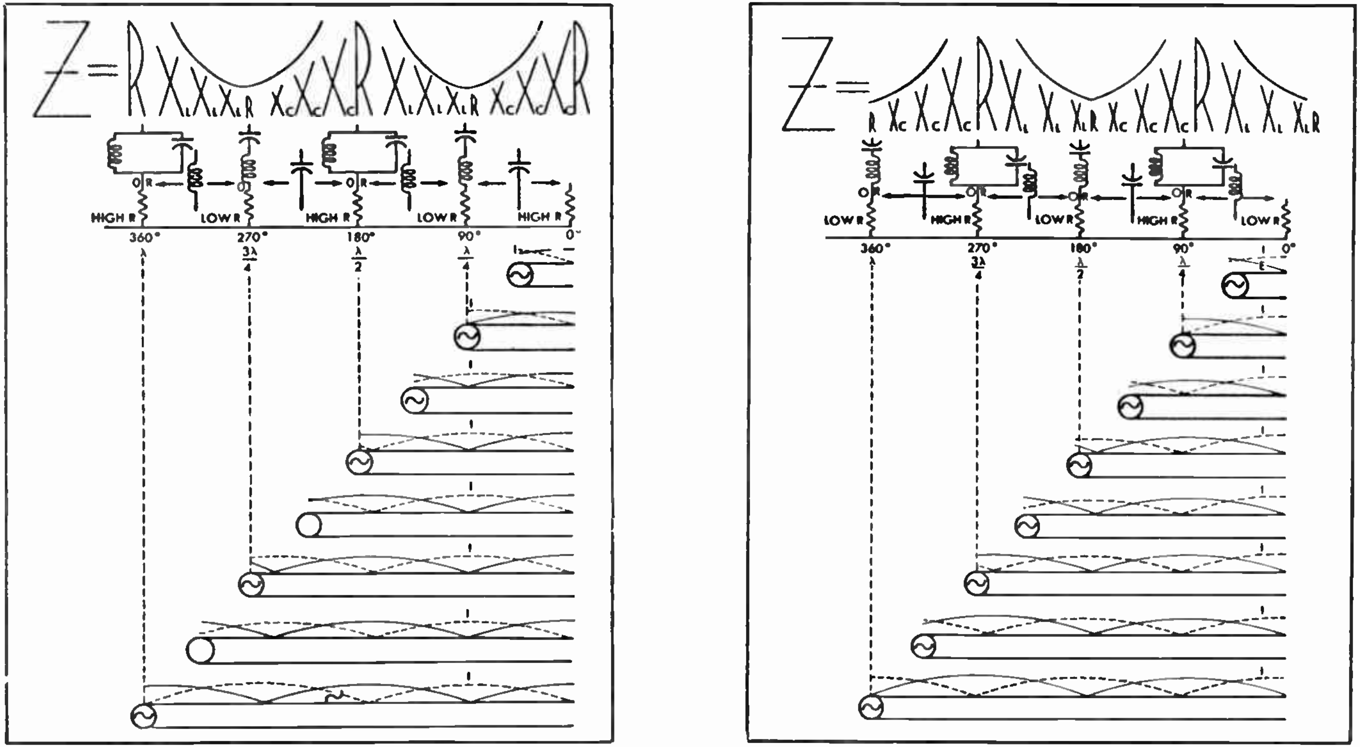


Fig. 2-20. Impedance charts.

little insight into velocity of propagation in a transmission line. The velocity of propagation of an electromagnetic wave in space is not just an arbitrary velocity, but is related in a definite way to the electric and magnetic properties of space. The velocity is given by

$$c = \frac{1}{\mu_0 \epsilon_0}$$

where μ_0 = magnetic permeability of free space

ϵ_0 = permittivity of free space

As stated earlier, in the rationalized mks system of units, these properties of space are given by

$\mu_0 = 1.26 \times 10^{-6}$

$\epsilon_0 = 8.85 \times 10^{-12}$

Thus

$c = 1.26 \times 8.85 \times 10^{-18} / 1 = 300 \times 10^{6}$ meters per second

You will probably never have any occasion to use the values of μ_0 and ϵ_0, but the above equation will shed a little more light on the subject of propagation.

We can write a somewhat similar equation for the velocity of propagation in a transmission line:

$$v = \frac{1}{\sqrt{LC}}$$

where L is the inductance of the line in henries per unit length, and C is the capacitance of the line in farads per unit length. The unit of length may be anything—feet, meters, etc.—just as long as the same unit is used for both inductance and capacitance.

The relationship between this equation and the preceding one for the velocity of propagation in free space can be appreciated by noting that permeability enters into the value of inductance, and permittivity enters into the value of capacitance. In an open-wire line in which the wire is made of nonmagnetic material and most of the space between the wires is filled with air, the velocity of propagation will be very close to that in free space. If, however, the space between the conductors is filled with a material that has a dielectric constant greater than 1, the capacitance will increase and the velocity of propagation will decrease. In the extreme case of a coaxial line that is filled with a dielectric material such as polyethylene, the velocity of propagation is as low as 60% of the value of propagation in free space.

Manufacturers of transmission lines specify the velocity of propagation in their products in terms of a *velocity factor* (VF). The velocity factor is the ratio of the velocity in the line to the velocity in free space. It is expressed either as a decimal or a percentage. Thus, for example, if a line has a velocity factor of 0.8, or 80%, the velocity of propagation will be 80% of the velocity in free space.

Now we are in a position to find wavelengths in actual transmission lines. The physical length λ of a wavelength in meters in a particular type of transmission line is given by

$$\lambda = \frac{300}{f} \times \text{VF}$$

where VF is the velocity factor of the line and f is the frequency in megahertz.

The number of wavelengths in a given physical length of transmission line is given by

$$l_{\lambda} = \frac{l}{\lambda}$$

where l_{λ} = number of wavelengths
l = physical length of the line in meters
λ = length of a wavelength in the particular cable as given by the preceding equation

In many applications it is more convenient to express the electrical length of a transmission line in electrical degrees. Inasmuch as there are 360° in a wavelength, the length of a cable in degrees is given by

$$\lambda_{\lambda}^{\circ} = \lambda / l \times 360 \times \text{VF}$$

where the symbols have the same meaning as in the preceding equations.

LOSSES IN TRANSMISSION LINES

So far, in all of our discussions of transmission lines, we have considered only ideal lines with no losses of any kind. In many practical problems we can take this approach and ignore losses. In other cases, losses must be considered.

In any practical transmission line there are two kinds of losses—those that result from the series resistance of the conductors in the line and those that result from leakage between the conductors of the line. These two types of losses can be taken into consideration by adding two components to our equivalent circuit for a transmission line (Fig. 2-21). The series resistance *R* represents losses due to the resistance of the conductors, and the shunt component *G* represents losses due to leakage between the conductors. The shunt component is more conveniently considered as a conductance; that's why we use the symbol *G*. In most broadcast applications the leakage is so low that *G* is very low and can be ignored. (Remember, a low conductance corresponds to a high resistance.)

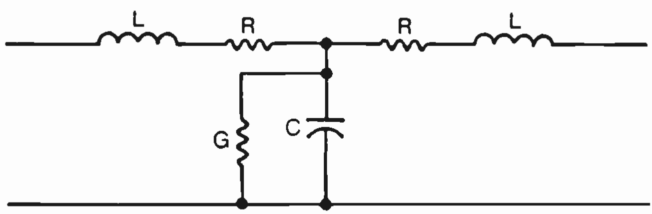


Fig. 2-21. Loss resistances in transmission lines.

The loss that is significant in broadcast applications is the series resistance of the conductors. The most common type of transmission line used in broadcasting is the coaxial line, so we use this type of line in our investigation of losses. Because the inner conductor is smaller than the outer one, its resistance is higher and it accounts for most of the loss in a coaxial cable.

Due to the *skin effect*, RF currents only flow in the outer skin of a conductor. The skin depth decreases as frequency increases, hence the resistance increases. As a matter of fact, the series resistance of a coaxial cable increases very nearly as the square root of the frequency. Thus the loss in a coaxial cable will be much greater at FM and TV frequencies than in the standard broadcast band.

It is interesting to note that there is a relationship between the characteristic impedance of a coaxial cable and the loss that it introduces. Assume that the diameter of a coaxial cable is some known value. It might appear that we could reduce the loss by increasing the diameter of the *inner* conductor. To some extent this is true, but as the diameter of the inner conductor is increased, the characteristic impedance of the cable is lowered. This means that the voltage-to-current ratio becomes smaller. In other words, more current will be required to transmit a given amount of power. Since the loss is proportional to the square of the current, we will eventually reach a point where losses actually increase as the diameter of the inner conductor increases.

Suppose that we take the opposite approach and *reduce* the diameter of the inner conductor. This raises the characteristic impedance of the cable and reduces the current, but it also raises the resistance of the inner conductor, so a point will again be reached where the losses increase.

Thus there is an optimum ratio of the outer- and inner-conductor diameters that will result in minimum loss. In cable where the space between the conductors is air, this ratio is about 3:1, giving a characteristic impedance of about 70 ohms. This value is rarely used in broadcasting because of other considerations, such as power-handling capacity, which optimizes at a value of characteristic impedance closer to 50 ohms.

There is another effect of losses in transmission lines that is rather unexpected. When we take the two resistances in Fig. 2-23 into consideration, the equation for the characteristic impedance becomes

$$Z_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}}$$

where L and C are the inductance and capacitance per unit length, R and G are resistance and conductance per unit length, and ω is $2\pi f$. From this equation we can see that the equation used earlier results when R and G are small enough to be ignored. Surprisingly, if the series resistance R becomes large, the characteristic impedance will have a reactive component. In broadcast work, lines are short and the $j\omega L$ term of the equation is high, so the reactive portion of the characteristic impedance is negligible.

REFLECTION COEFFICIENT

If we know the characteristic impedance of a transmission line, its length, and the load impedance connected to it, we have enough information to compute the voltage and current at any point on the line, as well as the standing-wave ratio. There are many different ways in which the parameters of a transmission line can be manipulated mathematically. In the process many different characteristics of transmission lines are described. The practical value of this is that we have many different measurements we can make to obtain the information we need.

At standard broadcast frequencies we make heavy use of impedance bridges, so we must be able to determine the behavior of a transmission line from impedance values. At the higher frequencies used for FM and TV broadcasting, we usually use some sort of *reflectometer*, which tells us the standing-wave ratio, or the forward and reflected power on the transmission line. Therefore we must be able to use these parameters to determine the behavior of the transmission lines. By means of several equations we can usually find what we need to know about this behavior from the information that is available from our instruments.

At the load of a transmission line, and all along the line for that matter, we consider three different voltages. The first is the forward voltage V_f, which travels down the line toward the load. The second is the reflected voltage V_r, which travels back along the line toward the source whenever the load impedance is not equal to the characteristic impedance of the line. The third voltage interest is the actual voltage along the line, which is the vector sum of the forward and reflected voltages at each point along the line. This voltage, as we have seen, varies with the distance along the line and is called a *standing wave*.

One useful parameter that specifies the nature of the reflection is the *reflection coefficient* K. It is a vector and is the ratio of the reflected voltage to the forward voltage, which is given as

$$K = \frac{V_r}{V_f}$$

The usual way of specifying the reflection coefficient is in polar form: $K \angle \theta$.

The reflection coefficient is merely another way of specifying what we have already specified in other terms. It is not surprising, therefore, to find that the reflection coefficient is related to the load impedance Z_L and the characteristic impedance Z_0 of the line. The relationship is

$$K = \frac{Z_L - Z_0}{Z_L + Z_0}$$

This equation can be rearranged to the following form, which is useful in some applications.

$$K = \frac{Z_L}{Z_0} - 1 \Big/ \frac{Z_L}{Z_0} + 1$$

The reflection coefficient is related to the standing-wave ratio by the equations

$$\text{VSWR} = \frac{1 + |K|}{1 - |K|} \qquad |K| = \frac{\text{VSWR} - 1}{\text{VSWR} + 1}$$

The bars in $|K|$ mean that, in these two equations, we are only interested in the magnitude of the reflection coefficient; we don't need the phase angle, because a standing-wave ratio is not a vector. It tells us the ratio of the maximum to minimum voltage on a line, but it doesn't tell us where the maximum and minimum voltage occur along the line.

Two additional concepts that are useful for dealing with transmission lines are *forward power* and *reflected power*. These concepts can be extremely troublesome if not properly understood. We can avoid confusion by remembering that power is merely the rate of flow of energy; energy is the more fundamental concept. Bearing this in mind, we can define the *forward power* on a transmission line as the average rate at which energy moves from the source toward the load. If the line was terminated in its characteristic impedance, the forward power would be the same as the actual power delivered to the load, assuming that there are no losses in the line.

When a line is not terminated in its characteristic impedance, some of the energy will be reflected from the load. Thus we can define *reflected power* as the average rate at which energy flows back from the load, along the line toward the source.

There are several points to keep clear about forward and reflected power. They are both merely ways of expressing the rate at which energy flows back and forth along a transmission line when the load impedance is not equal to the characteristic impedance of the line. This rate has little to do with how much power is being delivered by the transmitter. For example, it is possible to have a transmission line with a forward power of 150W and a reflected power of 50W with a transmitter delivering only 100W. The forward and reflected power deal only with the energy flow on the line that results from the line being mismatched. A good example of this is the ideal lossless line that is open at the receiving end. If the characteristic impedance of the line was 50 ohms and the voltage applied to the line by the transmitter was 100V, both the forward and reflected powers would be 200W, and yet, once the standing wave was set up, the transmitter wouldn't be delivering any power at all.

Obviously, if a transmission line is terminated in its characteristic impedance, there will be no reflected power and no standing waves. If the line is mismatched, there will be standing waves, and the reflected power will no longer be zero. Both standing waves and reflected power are measures of the same thing. We can convert from forward and reflected power to standing-wave ratio by the equation

$$\text{VSWR} = \frac{1 + \sqrt{\dfrac{\text{reflected power}}{\text{forward power}}}}{1 - \sqrt{\dfrac{\text{reflected power}}{\text{forward power}}}}$$

PRACTICAL TRANSMISSION LINES

There are three types of transmission lines that are used in broadcasting: open-wire lines, coaxial cables, and hollow waveguides. The coaxial cable has almost completely replaced the open-wire line, but there are still a few AM stations that use open-wire lines. Waveguides are only used at the UHF TV.

Open-Wire Line

The earliest transmission line used in broadcasting was the open 2-wire line, shown in Fig. 2-22. This line is simple and reliable, but unfortunately the fields from the conductors are strong at an appreciable distance from the line, with the result that any conductors in the vicinity of the line will disturb its characteristics. If a transmission line of this type is not to radiate energy, the currents in the two conductors must be equal in magnitude and opposite in direction. If any surrounding structure—or even the ground, for that matter—is closer to one conductor than the other, there will be capacitive currents that will cause the line currents to be unbalanced. Thus the line will radiate.

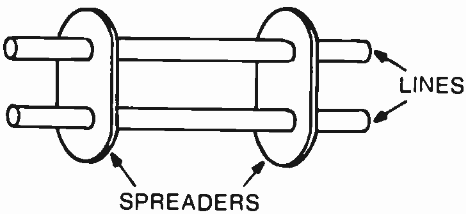


Fig. 2-22. Two-wire transmission line.

A somewhat more recent open-wire line uses five or six conductors in an arrangement such as that shown in Fig. 2-23. The outer conductors in this arrangement are all at the same potential and are connected to the grounded side of the transmitter and the antenna. This arrangement is something like a crude approximation to a coaxial cable. Unfortunately the conductor spacing between the spreaders will change as the wires swing in the wind, which will cause the characteristic impedance of the line to vary. These lines are rapidly being replaced by coaxial cable.

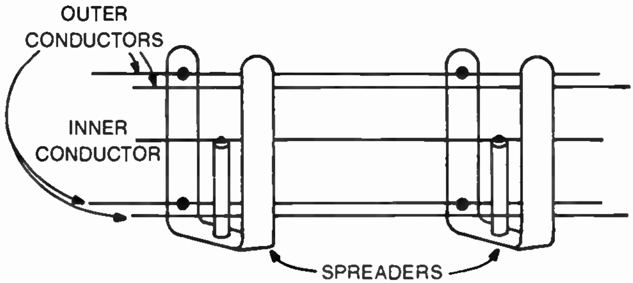


Fig. 2-23. Five-wire open transmission line.

Solid-Dielectric Coaxial Cable

Figure 2-24 shows a coaxial cable that consists of a solid or stranded inner conductor with a braided outer conductor. The space between the inner and outer conductors is filled with a solid dielectric material such as polyethylene. The entire assembly is covered with a weathertight plastic jacket. Usually the power-handling capacity of this type of line is quite limited, and it is rarely used except in some low-power AM stations. This cable is flexible, and for this reason has often been used in sampling systems for directional antennas. In many cases, however, it is being replaced because it tends to be unstable, expecially when the ambient temperature varies over a wide range.

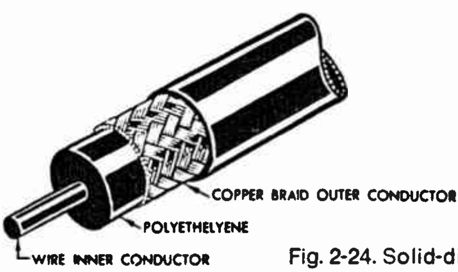


Fig. 2-24. Solid-dielectric coaxial cable.

Semirigid Coaxial Cable

Semirigid coaxial cable is made with soft-drawn copper inner and outer conductors. The line is not particularly flexible, but it can be bent a few times before breaking, and therefore it is easy to fit to a particular application. The cable is made by a continuous process and is shipped on reels. Thus it is possible to get single runs that are long enough to reach from the transmitter to the antenna in many stations. This avoids the necessity of making splices and joints, which are time consuming and potentially troublesome.

In the semirigid cable the inner and outer conductors are spaced by either beads or a helix of dielectric material. These lines are becoming popular for broadcast use because they are easy to install and because very low standing-wave ratios can be obtained. Figure 2-25 shows a sketch of a helical insulated line.

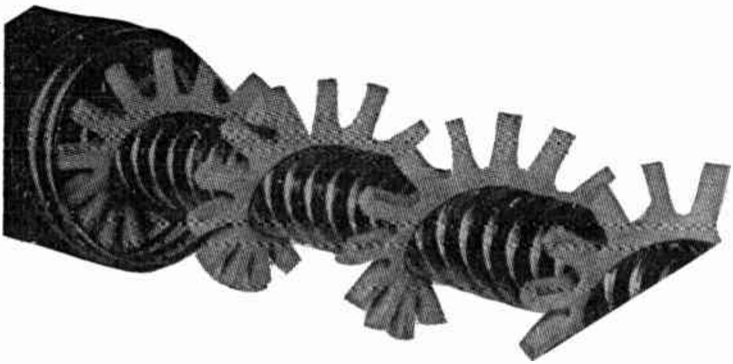


Fig. 2-25. Semirigid line with helical insulator.

Rigid Coaxial Lines

The coaxial line with the lowest losses and the highest power-handling capability is the rigid line, which comes in sizes of up to 6 in. diameter. Because it is rigid, it cannot be shipped on reels. It is usually supplied in 20-foot lengths. The lengths are fastened together by flanges and inner-conductor projections, called *bullets* (Fig. 2-26).

Many rigid coaxial lines are pressurized with nitrogen or dry air to keep moisture out of the space between the

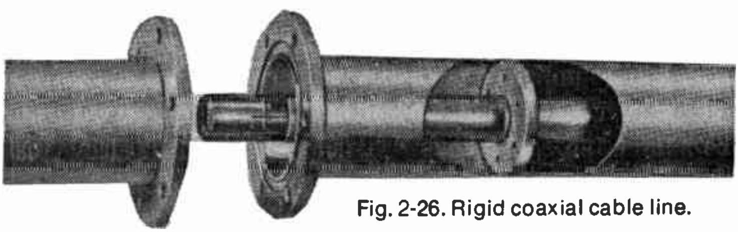


Fig. 2-26. Rigid coaxial cable line.

Table 2-1. Typical Attenuation Values for Coaxial Lines.

Line	Attenuation, dB per 100 ft. at the Following Frequencies, MHz						Velocity Factor
	1.0	5.0	10.0	50.0	100	500	
RG8/U	0.175	0.405	0.582	1.37	2.1	5.2	0.66
RG-17/U	0.061	0.158	0.238	0.62	0.95	2.7	0.66
Rigid 7/8 in.	0.0375	0.089	0.124	0.28	0.42	1.1	0.87
1 5/8 in.	0.0195	0.043	0.063	0.147	0.213	0.48	0.96
3 1/8 in.	0.0104	0.023	0.033	0.073	0.108	0.24	0.98
6 1/8 in.	0.0049	0.011	0.016	0.033	0.049	0.11	0.98
Heliax* 7/8 in.	0.034	0.077	0.11	0.25	0.36	0.46	0.92
3 in.	0.013	0.029	0.042	0.097	0.150	0.33	0.93
5 in.	0.0072	0.017	0.023	0.053	0.076	0.19	0.93

* Registered trade name

conductors. Pressure gauges are provided so that any leaks in the line can be detected.

The selection of a particular line is based on allowable loss or attenuation, power-handling capability, and ease of installation and maintenance. The attenuation of a line increases with frequency. Table 2-1 shows the attenuation of several different types of coaxial cables at various frequencies.

Chapter 3
Radiation and Propagation

There is no doubt that radiation is the least understood aspect of broadcasting. Most textbooks on the subject are of little help, because they are either too mathematical or too superficial. In this chapter we will review the subject of radiation, using no more mathematics than necessary. We will start out with the half-wave antenna because it is easiest to understand, then we will consider antennas that are more commonly used in broadcasting.

In considering radiation we will not be very anxious to make everything strictly rigorous, rather we will take an approach that will make the subject more palatable.

Before getting into the details of how an antenna radiates energy, we should make a distinction between two different types of fields we will encounter:

1. The *induction field* about an antenna is the same as the electric or magnetic field that is found around any conductor that carries electricity. Its intensity diminishes rapidly with distance, or to state it more rigorously: The intensity of the induction field varies inversely with the square of the distance from the source of the field. Thus, if we move twice as far from the source of an induction field, the intensity of the field will be one-fourth as great. This field diminishes so fast that it is of no value for broadcasting.

2. The other field around an antenna is the radiation field. This field is properly called an *electromagnetic* field because it has both electric and magnetic components. These components act together to propagate the signal. The radiation field diminishes much more slowly with distance than the induction field; its intensity is inversely proportional to the distance from the antenna. It is this characteristic of the radiation field that makes it useful for broadcasting.

As discussed in the preceding chapter, when electromagnetic energy is carried by a transmission line, it isn't carried in the wires, but in the electric and magnetic fields associated with the wires. The conductors merely serve to guide the energy and cancel the fields at some distance from the transmission line. The object of a transmission line is to get energy from one place to another with a minimum of radiation. The object of an antenna is to radiate a maximum amount of the energy fed to it.

INDUCTION FIELD

In Fig. 3-1A we have a source of RF energy connected to a transmission line. For the sake of simplicity, we will assume that it is a 2-wire open line. The receiving end of the line is open, so if it is an ideal line, all of the energy reaching the receiving end is reflected back toward the source. We can think of the last 1/4 wavelength of the line, marked in the figure, as an open-ended quarter-wave section of transmission line. Such a section looks electrically like a series-resonant circuit. If we bend the ends of the last 1/4 wavelength of the line out so that they are at right angles to the wires in the line (Fig. 3-1B) we have what is commonly called a *half-wave dipole* antenna at the receiving end of the line.

By analogy with an open quarter-wave section of transmission line, the half-wave dipole seems electrically like a series-resonant circuit when we look into its terminals. The half-wave dipole differs from an ordinary series-resonant circuit in many ways, but we can gain a little insight into antennas by comparing them.

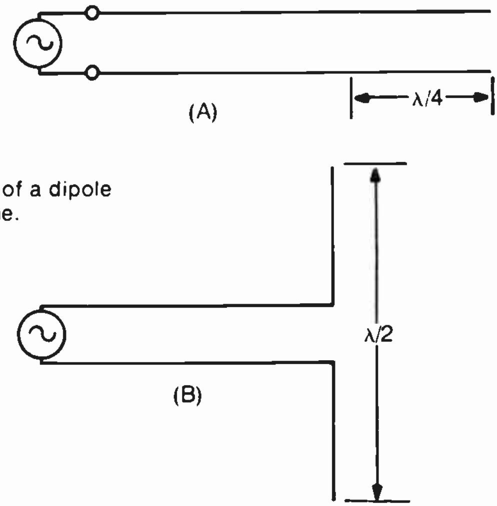


Fig. 3-1. Development of a dipole from a transmission line.

In Fig. 3-2A there is no charge in the capacitor in the series-resonant circuit. When the signal is applied, current rushes in to charge the capacitor. Because the current is maximum, the magnetic field around the coil is maximum. An analogous thing happens with the dipole. At the instant shown, there is no charge on the ends of the dipole. The current is maximum, carrying positive charges to one end and negative charges to the other. Thus the magnetic field around the dipole is maximum.

Figure 3-2B shows the situation a quarter of a cycle later. In the resonant circuit the capacitor is fully charged, so the current, and hence the magnetic field, is zero. All of the energy is stored in the electric field of the capacitor. Back at the dipole the ends are charged and the current and the magnetic field are zero. The electric field between the ends of the dipole is maximum.

In Fig. 3-2C we have progressed still another quarter of a cycle. This is one-half cycle, or 180°, later than when we started so we can expect things to be 180° out of phase with what they were in Fig. 3-2A. The current in the resonant circuit and dipole is maximum, as is the magnetic field, but the current is flowing in the opposite direction of the current in Fig. 3-2A. It is an easy step to Fig. 3-2C, where the charge is maximum and the current and magnetic field are minimum.

Thus in Fig. 3-2 we see that the energy in both a series-resonant circuit and a half-wave dipole is stored alternately in the electric and magnetic fields. Assuming that there are no losses and no radiation, the energy simply pulsates back and forth between the electric and magnetic fields around the antenna. These fields constitute what we called the *induction field* of the antenna. Note that they are 90° out of phase with each other.

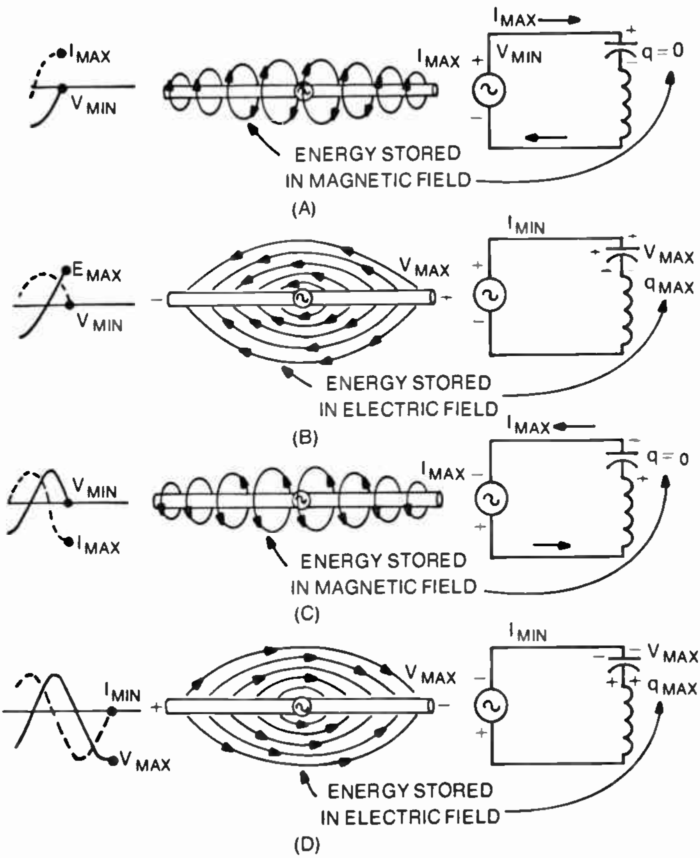


Fig. 3-2. Analogy of dipole and series-resonant circuit.

If this were all there is to it, antennas would be simple, and easy to understand. Of course, antennas must radiate energy. To understand radiation, we must go back to some of the most fundamental concepts of electricity and modify them slightly from the way that we learned them. Most of us first studied electricity from the viewpoint of circuit theory. This is fine when we are dealing with circuits, where currents and voltages stay in the wires where we want them. It is inadequate when we deal with antennas, where we don't want the signals confined to circuits, but radiated through space.

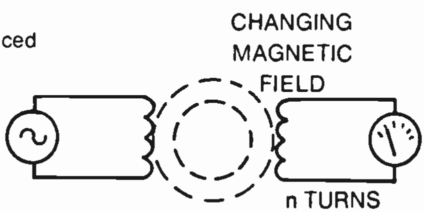
INDUCED VOLTAGE

Figure 3-3 shows the familiar principle of induced voltage. When a conductor intercepts a changing magnetic field, a voltage is induced in the conductor. The induced voltage is proportional to the *rate of change* of the magnetic field. The equation for the induced voltage is

$$V = -n \frac{\Delta\phi}{\Delta t}$$

where V is the induced voltage, n is the number of turns on the coil, $\Delta\phi$ is a small change in magnetic flux, and Δt is a small change in time. Thus $\Delta\phi/\Delta t$ represents the rate of change of the magnetic field. The minus sign indicates that the polarity of the induced voltage is such that any current it causes to flow produces a magnetic field opposing the changing field that caused the induced voltage. This is about as far as we ever carry this principle when we are studying circuit theory.

Fig. 3-3. Principle of induced voltage.



As far as the principle is concerned, the coil isn't necessary at all. Whenever we have a changing magnetic field, we have a changing electric field—that is, we have an induced voltage, even in an insulator. Of course, unless we have a conductor, we have no way of measuring the induced voltage.

The general principle, then, is that whenever we have a changing magnetic field, we also have a changing electric field, even in free space. The intensity of the electric field is proportional to the rate of change of the magnetic field. In

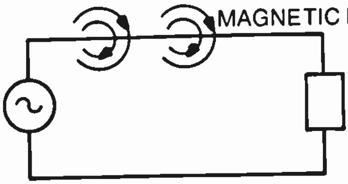


Fig. 3-4. Magnetic field induced by current.

fact, when we are dealing with radiation, it is best to think of the electromagnetic field as the fundamental concept, and the electric and magnetic fields as components of it.

ELECTRIC DISPLACEMENT

Whenever an electric current flows in a conductor, there will be a magnetic field surrounding the conductor (Fig. 3-4). Before we get into radiation, we must explore this concept in more detail. Figure 3-5A shows an RF source connected through leads to a capacitor. We know that an RF current will flow in the leads and that, as a result, a magnetic field will enclose the leads. Of course, the magnetic field will vary at the frequency of the source. Now let's look at what happens inside the capacitor.

We know that the dielectric material between the plates of the capacitor is actually an insulator and that no current, or at least no electrons, can pass through it. When we studied elementary electricity, we learned that there is an *apparant current* in the capacitor, which is actually the result of electrons "piling up" on one plate and draining off the other plate. For this reason the apparent current through a capacitor is also called a *displacement* current.

We shall now consider the interesting question of whether there is a magnetic field associated with the displacement current inside a capacitor. This isn't an easy question. In fact, there was a great deal of debate on the subject among early workers in electricity. If we think of an electric current only as a flow of electrons, we are tempted to say that displacement current inside a capacitor is a fictitious thing and couldn't possibly produce a magnetic field. This is wrong. James Clerk Maxwell, a Scottish physicist, was the first to postulate that there is in fact a magnetic field associated with displacement current. He used this assumption in deriving his now-famous equations.

Earlier we found that whenever we have a changing magnetic field, we also have a changing electric field. Now we add to this that whenever we have a changing electric field, even when we do not have a conductor (Fig. 3-5B), we also have a changing magnetic field. Thus we see that the two fields are inseparable. Inasmuch as we see that these fields can exist and produce each other even in free space, the concept of radiation is becoming clearer.

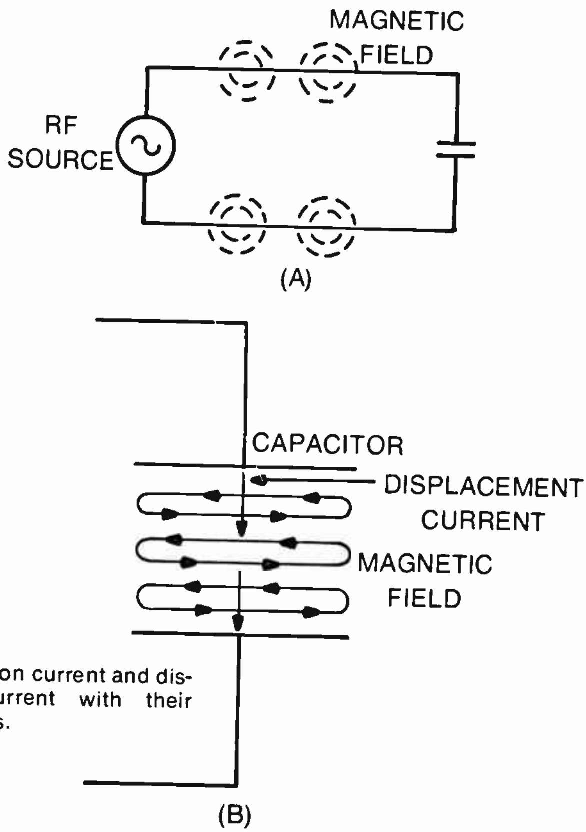


Fig. 3-5. Electron current and displacement current with their magnetic fields.

PROPAGATION TIME

Electromagnetic phenomena do not take place instantaneously; there is some time required. Electromagnetic fields do not travel at infinite velocity; they travel at the velocity of light, which, although fast compared with any other phenomenon, is still not infinite.

To see why we might need a time delay to explain radiation, look at Fig. 3-6. Here we have a half-wave dipole at the same instant of time as in Fig. 3-2B. The charge between the ends of the dipole is maximum, and the current is minimum. According to the law of charges, unlike charges attract. How, then, can the opposite charges flow away from each other to opposite ends of the dipole (Fig. 3-6)? If electromagnetic phenomena took place instantaneously, the charges could not do this. The fact is, the positive charges at the top of the dipole, at the instant shown, are actually still being repelled by the positive charges that were at the other end of the dipole a half-wave earlier in time. This occurs because it takes 1/2 wavelength (180°) of time for electromagnetic energy to travel 1/2 wavelength through space.

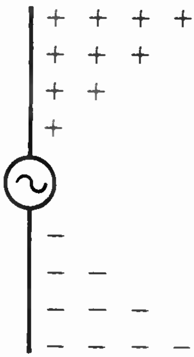


Fig. 3-6. Charges on dipole at one instant.

RADIATION

Once we have one kind of field, it will generate the other, and the action will continue causing the energy to propagate through space. This is shown in a rather crude fashion in Fig. 3-7. Here we have an electric field at the left, and as it

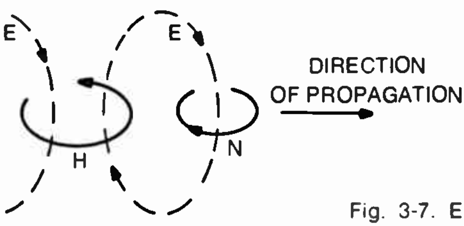


Fig. 3-7. Electric and magnetic fields.

collapses, it sets up a magnetic field, which in turn sets up another electric field, and so on.

Just how the field gets free of the antenna in the first place isn't as easy to visualize. A rough idea of what happens can be gained from Fig. 3-8. In Fig. 3-8A the charges at the ends of the antenna are maximum, as is the electric field. In Fig. 3-8B the current is such as to reduce the charges and hence the field. The lines of the field at the antenna are brought closer together. The field doesn't collapse completely, however, because some time is required for all electromagnetic effects to be observed at a distance. Thus, as the opposite charges on the ends of the antenna come together and cancel each other, the lines of the field become closed on each side of the antenna (Fig. 3-8C). About this time, the field in the dipole reverses, so it repels the electric lines that have become detached, causing them to propagate through space as shown in Fig. 3-8D. Thus we will have a field with both electric and magnetic components, moving through space at the speed of light.

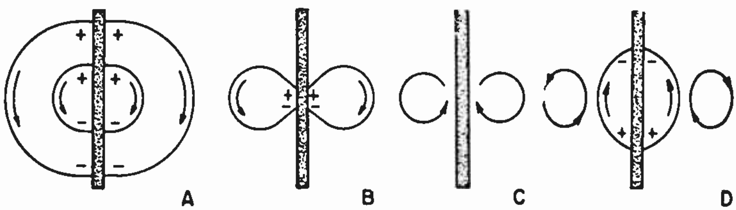


Fig. 3-8. Creation of closed electric lines at an antenna.

The electric and magnetic lines are closed paths, so near the antenna, the wavefront will appear to be spherical. However, the sphere becomes large rapidly, and as soon as we are some appreciable distance from the antenna, we can consider the wavefront to be a plane, with the electric and magnetic lines at right angles.

In antenna work the electric component of the field is usually called the *E-field*, and the magnetic component is called the *H-field*. This is because the letter *E* is used to symbolize the electric field intensity, and the letter *H* is used to symbolize magnetic field intensity.

POLARIZATION

If a straight antenna is used to radiate a wave, the electric field lines are parallel to the antenna. So far, we have only considered antennas in space; we have not considered the ground. In broadcast work, antennas are usually either vertical or horizontal with respect to the ground. The electric and magnetic fields are also so oriented. We call the direction of the fields the *polarization* of the wave. Traditionally the direction of the *electric* field, rather than the magnetic field, is taken as the reference. Thus the wave from a vertical antenna, whose electric field is perpendicular to the ground, is called a *vertically polarized* wave. Hence all waves encountered in the standard broadcast band are vertically polarized. Likewise, the wave from a horizontal antenna, whose electric field is horizontal, is called a *horizontally polarized* wave. Waves from FM and TV antennas are usually horizontally polarized.

It is possible to have a combination of vertical and horizontal polarization, in which the electric field actually rotates with respect to the ground. This is called *circular* polarization. At this writing, it is being used extensively in FM broadcasting and experimentally in TV broadcasting.

Figure 3-9 shows the orientation of the fields for vertical and horizontal polarization. It shows the electric field lines at an instant of time. The lines are actually propagating through space in the direction shown (left to right). The shaded sine waves in the figure show the relative field intensity at various points in space at a particular instant.

FIELD INTENSITY

The intensity of a radiated wave is measured in volts per unit of distance. The fundamental unit is the *volt per meter* (*V*/m). A field having an intensity of one volt per meter is very strong compared with most of those encountered in broadcasting, so the *millivolt per meter*, or even *microvolt per meter*, is commonly used.

A field having an intensity of one volt per meter (1 V/m) would induce one volt in a conductor one meter long if the conductor was held parallel to the electric field and perpendicular to the direction of the wave. The radiation field

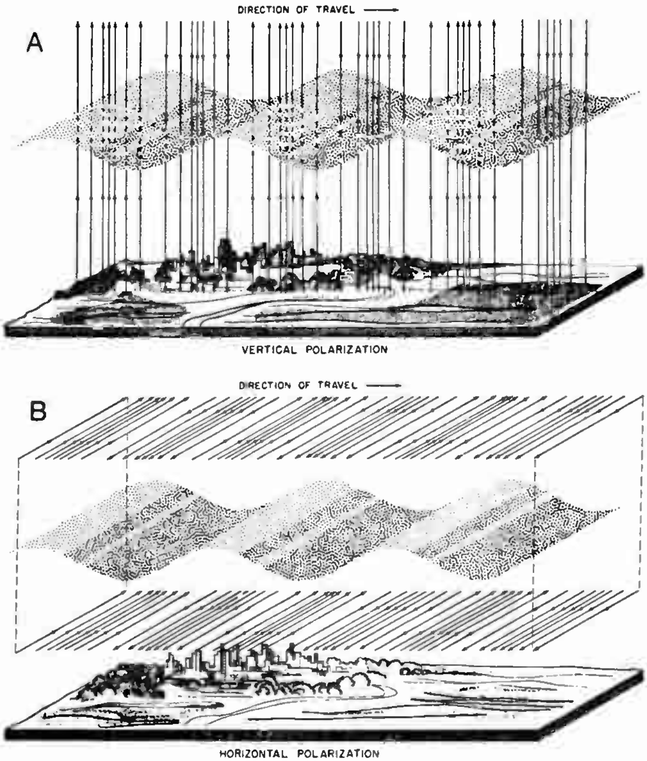


Fig. 3-9. Vertical and horizontal polorization.

varies inversely with the distance from the antenna. Thus if a wave has a field intensity of 50 mV/m at a point one mile from an antenna, it will have a field intensity of 25 mV/m at a distance of 2 miles from the antenna.

The intensity of an electromagnetic field in volts per meter is really only a measure of the electric field or the electric component of the electromagnetic field. Fortunately we need not specify the intensity of both the electric and magnetic components of the electromagnetic field to describe it completely. Once we are far enough away from the antenna that we can consider the waves to be plane waves, there is a very definite relationship between the electric component of a wave and the magnetic component. We only need to specify one of them to completely describe the intensity of the wave, and the electric component is traditionally used for this purpose.

There is an intriguing relationship between the magnitudes of the electric and magnetic components of a plane wave in free space. The ratio of the two is

$$\frac{E}{H} = \frac{1}{\sqrt{\mu_0 / \epsilon_0}}$$

where E is the intensity of the electric field in volts per meter, H is the intensity of the magnetic component in amperes per meter ϵ_0 is the permittivity, and μ_0 is the permeability of free space. Thus we have

$$\frac{E}{H} = \frac{1}{\sqrt{\dfrac{8.85 \times 10^{-12}}{1.26 \times 10^{-6}}}} = 377 \quad \frac{\text{V/m}}{\text{A/m}}$$

Since the meters in the numerator and denominator cancel out, we have

$$\frac{E}{H} = 377\text{V/A}$$

That is, the ratio of the electric component to the magnetic component of a plane wave is 377V per ampere. Now, we know that the ratio of volts to amperes is impedance, so we conclude that free space has a *characteristic impedance*, or *wave impedance*, of 377 ohms. This numerical value is not of particular interest, but it does make it clear why we have to measure or specify only one component of the field.

FIELD INTENSITY VERSUS DISTANCE

The radiation intensity of a wave varies inversely with the distance from the antenna. The reason for this will become clear if we go back to the more fundamental concepts of power and energy. What the antenna is actually radiating is electromagnetic energy. It is this energy that causes the charges in receiving antennas to move when impinged on by the radiated wave. The rate at which energy is propagated by a wave can be specified in terms of watts per square meter.

Suppose that the antenna in Fig. 3-10 is radiating energy through the beam shown at a rate of 8W. At point *A*, the beam intercepts an area of just one square meter. We can say that the *power density*, or *wave power*, at that point is 8 W/m².

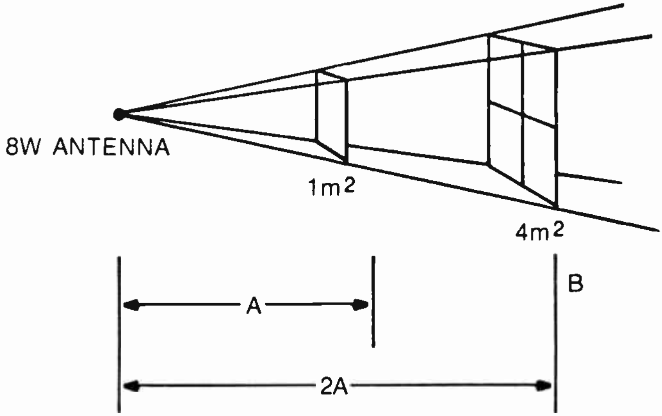


Fig. 3-10. Variation in power density with distance.

At point *B*, which is twice as far from the antenna as point *A*, the beam intercepts an area of 4 m². The beam still has a power of 8W, so the power density at point *B* is 2 W/m² (8/4 = 2). Thus, when we double the distance from the antenna, we decrease the power density to one-fouth. We can say, therefore, that the power density in watts per square meter *varies inversely with the square of the distance* from the antenna.

In standard broadcast antennas we are not interested in power density. Rather we measure signal strength in terms of field intensity in volts per meter. The field intensity is proportional to the square root of power density; therefore, the *field intensity varies inversely with the distance, not the square of the distance, from the antenna.*

GAIN AND DIRECTIVITY

Much antenna-design work involves getting an antenna to radiate more energy into one region than into another. One measure of this property of an antenna is its *gain*. This is a relative term. If we say that an antenna has a certain amount of gain, we must state some reference. We might, for instance, say that an antenna has a gain in a given direction of 2, referred to a half-wave dipole. This means that the antenna radiates twice as much power into a region as a half-wave dipole for the same value of transmitted power.

One reference that is widely used in antenna work is the *isotropic antenna*, which radiates uniformly in all directions. Of course, it is not possible to actually build an isotropic antenna. No real antenna will radiate equally in all directions. Nevertheless, the isotropic antenna is a useful reference because its performance is easy to calculate. Since it is a fictional device, we avoid the problem of making a standard antenna that will perform properly.

The field intensity around an antenna is always specified at some distance from the antenna. In broadcast work this distance is almost universally chosen as one mile. The formula for the area of a sphere is

$$A = 4\pi r^2$$

where r is the radius of the sphere. Thus if we have an isotropic antenna at the center of a sphere having a radius of one mile, and if the power radiated was 1 kW, the power density at the surface of the sphere would be

$$p = \frac{1000}{4\pi(1609)^2} = 0.000031\ \mathrm{W/m^2} = 31\ \mu\mathrm{W/m^2}$$

Note: There are 1609m in a mile.

Since the wave impedance of free space is 377 ohms, we have a means of computing the field intensity when we know the power density. Using a formula that is analogous to the power formula

$$\text{Voltage} = \sqrt{pR}$$

we can say that the field intensity in volts per meter is given by

$$E = \sqrt{\frac{\text{watts}}{\text{meter}^2}} \times 377\ \text{ohms}$$

$$= \sqrt{0.000031} \times 377 = 0.1076\ \mathrm{V/m}$$

Thus the field intensity from an isotropic antenna at a distance of one mile with a radiated power of 1 kW is 107.6 mV/m.

The gain of an antenna is usually expressed as a *power* gain. If an antenna has a gain of 2, it will radiate twice as many watts per square meter into a given region as an isotropic antenna would. If we want to know the field intensity at one mile, produced by an antenna that has a gain of 2, with a radiated power of 1 kW, we must take the square root of the power gain. Thus

$$E = 107.6 \times \sqrt{2} = 152.2 \text{ mV/m}$$

Antenna gain is frequently expressed in decibels. The formula for computing the gain in decibels is

$$G_p = 10 \log p_{ref} / p$$

where p is the power density from the antenna under consideration, and p_{ref} is the power density from the reference antenna. The same units must be used for p and p_{ref}. In terms of field intensity the gain in decibels is given by

$$G_p = 10 \log \frac{E^2}{E_{ref}^{\ 2}} = 20 \log \frac{E}{E_{ref}}$$

where E is the field intensity of the antenna under consideration, and E_{ref} is the field intensity from the reference antenna. The same units must be used for E and E_{ref}.

DETERMINING PATTERN SHAPE

The statement that an antenna has gain suggests that we can get something for nothing. Of course, this isn't true. If an antenna radiates more energy into a region than would be radiated by an isotropic antenna, it must radiate less energy into some other region. The measure of how much energy an antenna radiates into various regions is called the *radiation pattern* of the antenna.

The mathematical procedure for computing the radiation pattern for an antenna of arbitrary shape and size is very involved. Often the equations cannot be solved, because their solution depends on an accurate knowledge of the distribution of charge and current all along the antenna. If the pattern of an antenna is known through measurements, the mathematician can go back and manipulate the equations until they agree with the measured results. In all but the simplest cases, he cannot accurately predict the pattern in advance.

Fortunately for the broadcast engineer, it isn't necessary to compute the patterns of basic antenna elements. This has been done many times, and the results are readily available. It is helpful, however, to know the principles involved, because they give some insight into the behavior of actual antennas.

Figure 3-11 shows an elementary antenna that is very short. Mathematically speaking, its length is infinitesimal. Since the antenna is very short, we can assume that the current is the same all through it. Of course, such an antenna is impossible to build, and it wouldn't be worth trying to approximate, because its losses would be extremely high. It is useful to consider, however, because a practical antenna can be thought of as being made up of a very large number of these elementary antennas—which are usually called *elementary dipoles*—and the contributions of all of them can be added to find the field from the real antenna.

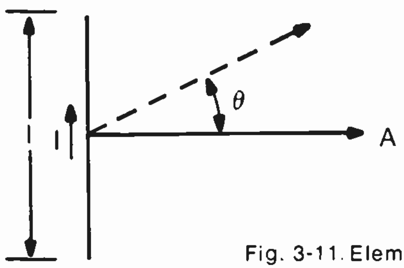


Fig. 3-11. Elementary infinitesimal dipole.

By just looking at the antenna of Fig. 3-11, we can predict a few things about it. From an earlier discussion we know that the electric field is parallel to the antenna, and the magnetic field is at right angles to the electric field. Thus the maximum radiation will be broadside to the antenna, toward point *A*. The radiation will be zero off the ends of the antenna. It isn't surprising, then, that the radiation falls off as the cosine of the angle θ. What we have no way of knowing—without solving some rather unpleasant equations—is the actual intensity of the field. Avoiding the unpleasant equations, we find that the field intensity is given by

$$E = \frac{60\pi}{d\lambda} Il \cos\theta$$

where E = field intensity in volts per meter
d = distance from antenna in meters
λ = wavelength of signal in meters
I = current in amperes
l = length of antenna in meters
θ = angle from a plane perpendicular to antenna

Strictly speaking, l should be infinitely small, and we should use calculus to find the field intensity. Actually, we can get a valid solution if we assume that l is small but finite. We can simplify the equation a little by substituting 1609m for d and moving λ under l. The equation then becomes

$$E = 0.117I\frac{l}{\lambda}|\cos|\theta$$

where l/λ is the length of our elementary dipole, expressed as a fraction of a wavelength. Thus, if the dipole is one electrical degree in length, and the current is one ampere, the field intensity becomes

$$E = 0.117 \times 1 \times \frac{1}{360} = 0.000325 \text{ V/m}$$

Thus the elementary dipole produces a field intensity of 0.325 mV/m one mile away, along the line broadside to the antenna, when the current is one ampere.

At the moment we have an expression for field intensity as a function of the current in the antenna, but we have no way to relate this to the actual power transmitted. The relationship involves the resistance seen by the current in the antenna. Our elementary dipole offers a resistance of 0.0061 ohms to the current flowing in it. Using Ohm's law, we can find that the current required for an elementary dipole to radiate 1 kW is

$$I = \frac{p}{R} = \frac{1000}{0.0061} = 405\text{A}$$

Substituting this current into our equation for field intensity gives

$$E = 0.325 \times 405 = 131.5\,\text{mV/m}$$

This equation shows that the field intensity from an elementary dipole has a maximum value of 131.5 mV/m. The field intensity varies with the cosine of the angle θ from a line that is broadside to the axis of the antenna (Fig. 3-12).

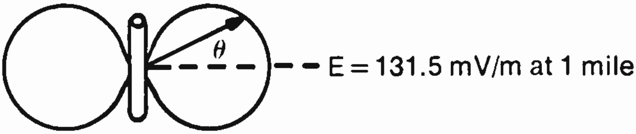


Fig. 3-12. Pattern of elementary dipole for 1 kW of radiated power.

We can use the mathematical tools that we have developed so far to find the pattern of a half-wave dipole. To do this we must assume that the half-wave dipole, which is 180° in length, is made up of 180 elementary dipoles. We can't simply multiply the results that we obtained earlier by 180, because the current is not uniform in the half-wave dipole. For a first approximation, we can assume that the current on the half-wave dipole is sinusoidal, being maximum at the center and zero at each end (Fig. 3-13). We can then break up the half-wave dipole into 180 elementary dipoles, each carrying the proper amount of current, and compute the contribution of each elementary dipole to the field intensity. We can then add these contributions together to find the field resulting from the half-wave dipole. The computations are rather lengthy, and we need not bother with the details here.

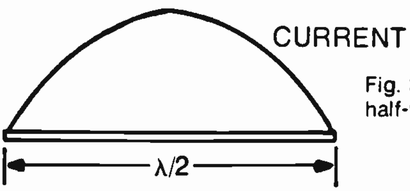
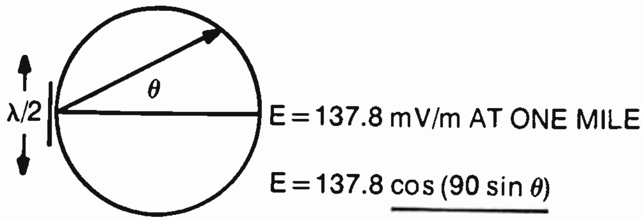


Fig. 3-13. Current distribution on half-wave dipole.

Figure 3-14 shows the radiation pattern of a half-wave dipole. It shows that the maximum field intensity at one mile is 137.8 mV/m for 1 kW of radiated power. Note that this and other theoretical patterns depend on the current distribution along the antenna itself. This point is important because there are many factors that influence the current distribution in practical antennas. Things such as lines that furnish current for tower lighting, other conducting structures in the vicinity, and even the guy wires on a tower will to some extent affect the current distribution. This effect, in turn, will result in some deviation of the actual radiation pattern from the computed theoretical pattern.



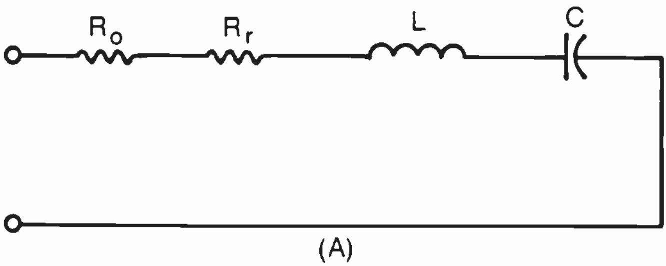
$$E = 137.8 \frac{\cos(90 \sin\theta)}{\cos\theta}$$

Fig. 3-14. Pattern of half-wave dipole for 1 kW radiated power.

ANTENNA IMPEDANCE

So far we have considered a current flowing in an antenna without regard as to how it happened to get there. Obviously, if we are to have a current in an antenna, we must feed energy to the antenna at some point. At this point, wherever it might be, we see an impedance. Inasmuch as energy enters the antenna and doesn't return, the impedance must have a resistive component. In addition to the energy that is radiated by the antenna, some energy is stored in the electric and magnetic fields in the near zone of the antenna. This means that the antenna impedance will also have inductive and capacitive reactive components. The actual amount of resistance and reactance seen looking into an antenna depends on what part of the antenna we feed, the physical dimensions of the antenna, and the frequency of operation.

Figure 3-15A shows a circuit that is a rough equivalent of a half-wave dipole. At the frequency at which the dipole is electrically 1/2 wavelength long, the inductive and capacitive reactances are equal and cancel each other. The equivalent circuit then becomes that shown in Fig. 3-15B, which consists merely of two resistances. One of these resistances (R_o) represents the *loss* or *ohmic* resistance of the antenna. Current flowing in this resistance is dissipated in the form of heat and is not available for radiation. The other resistance (R_r) is called the *radiation* resistance of the antenna. As far as we can see, looking into the terminals of the antenna, the energy that is radiated is dissipated in this resistance.



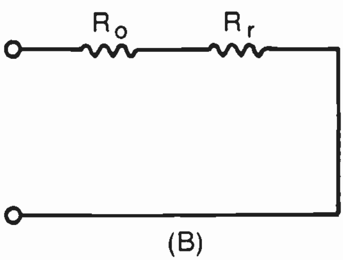


Fig. 3-15. Equivalent circuit of a half-wave dipole.

The total power entering the antenna is given by I^2R, where $R = R_r + R_o$. The purpose of the antenna is to release as much energy as possible through radiation and as little energy as possible through losses. For this reason the ohmic resistance R_o should be kept as low as possible. There is usually a minimum value below which it is impractical to reduce R_o. Therefore it is desirable to keep the radiation resistance high compared with the value of R_o. In general, short antennas tend to have low values of radiation resistance and, hence, high losses. The half-wave dipole has a radiation resistance of about 73 ohms, looking into its center.

The voltage and current are not constant throughout the length of an antenna. In the half-wave dipole the current is zero at the ends because there is no place for it to go, and it is maximum at the center. The voltage is just the opposite, being maximum at the ends of the antenna and minimum at the center. Impedance varies inversely with voltage, so it will be minimum at the center and maximum at the ends. The numerical value of the impedance varies from about 73 ohms to about 2500 ohms at the ends. Theoretically, the impedance at the ends would be infinite if the current actually went to zero. There is, however, always some capacitive current at the end of the antenna (*end effect.*)

At FM and TV frequencies it is often convenient to adjust the lengths of antenna elements so that they are resonant. This is usually impractical at standard broadcast frequencies, so the input impedance of standard broadcast antennas almost always has a reactive component.

In all broadcast services it is important that the impedance of the antenna not change significantly over the bandwidth of the signal. Because of the wider bandwidths involved, this consideration is most important in TV broadcasting.

VELOCITY OF PROPAGATION AND ANTENNA LENGTH

The velocity of propagation of electromagnetic waves in free space is very nearly 300,000,000 meters per second. If the velocity of a wave on antenna were the same as in free space, the wavelength on an antenna would be the same as in free space and would be given by

$$\lambda = \frac{300{,}000{,}000}{f}$$

where f is the frequency in hertz, and λ is the wavelength in meters. Like a transmission line, an antenna has inductance and capacitance, and these tend to retard the velocity of propagation. The larger the diameter of the antenna, the more capacitance per unit length. Thus the velocity will be lower in an antenna of large diameter than in a thin wire. Figure 3-16 shows the amount that the velocity of propagation is reduced as a function of the circumference in wavelengths.

The fact that 1/2 wavelength is shorter on an antenna than in free space causes a great deal of confusion. At standard broadcast frequencies we ignore the actual wavelength on the antenna and measure the height in electrical degrees, using the velocity of propagation in free space as a reference. (The reason for doing this is shown in Chapter 6.) On the other hand, FM and TV antennas that are 1/2 wavelength long take the velocity of propagation into consideration. With transmission lines, we also specify length in electrical degrees, but here we do take the velocity of propagation along the line into consideration.

NEAR AND FAR ZONES

An antenna has an induction field, which is useless as far as broadcasting is concerned, as well as the radiation field, which is what we use in broadcasting. Although we have no interest in the induction field, we must remember that close to the antenna the induction field is much larger than the

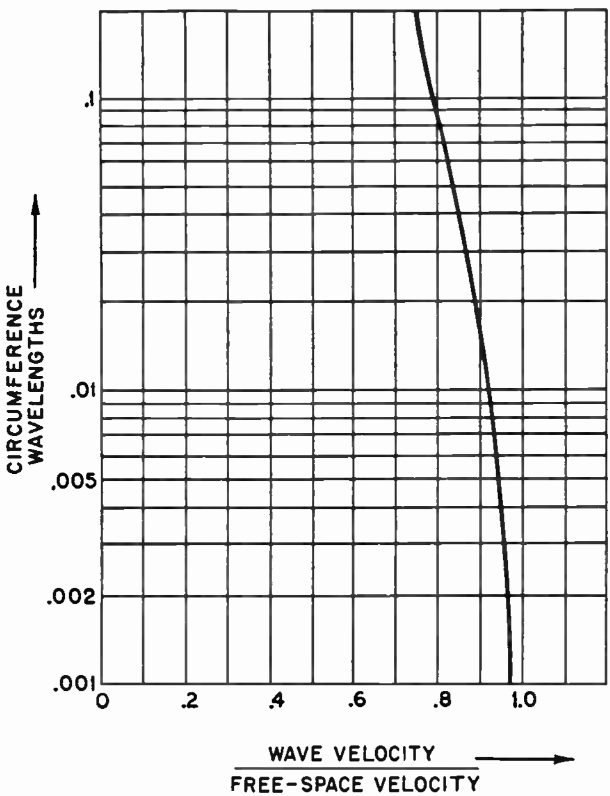


Fig. 3-16. Effect of antenna circumference on wave velocity.

radiation field. This means that any measurements that we might make on the radiation field must be made far enough away from the antenna that the induction field will not introduce errors.

It is customary to divide the region around an antenna into two zones—the *Fresnel* or *near* zone and the *Fraunhoffer* or *far* zone. The dividing line between the two zones is at a distance of $D^2/2\lambda$, where D is the largest dimension of the antenna, and λ is the wavelength, both in the same units. At this distance the induction and radiation fields are equal; beyond this distance the induction field diminishes with distance much more rapidly than the radiation field.

For standard broadcast antennas there is another consideration that limits how close to the antenna site we may take meaningful measurements. This is because many such antennas consist of several towers. In our field calculations we usually consider an antenna as acting as a point source. To make meaningful measurements, we must be far enough away from the antenna that it will look electrically like a point source. This distance is often much greater than the distance to the far field.

Chapter 4
Smith Charts

One of the most useful tools for solving antenna and transmission-line problems is the Smith chart, shown in Fig. 4-1. It can be used to find the standing-wave ratio, reflection coefficient, and impedances at various points in a feeder system with a minimum of mathematical calculation. In spite of its utility the Smith chart is not widely used by broadcast engineers.

The reason the Smith chart is not more popular is probably that it looks very complicated. This may be partly because the scales are circular rather than straight. The fact is, once one becomes familiar with the various scales, the Smith chart is no more difficult to use than any other graph, and it saves a considerable amount of labor. As we see, there are many advantages to using circular scales, not the least of which is that any value of impedance can be within the boundary of the graph.

Smith charts are available from most college book stores. They are available as either paper graphs or as plastic calculators with movable scales. The calculator is handy in applications where many different problems are to be solved, but the paper graph is fully adequate for broadcast applications and provides a permanent record of the computations.

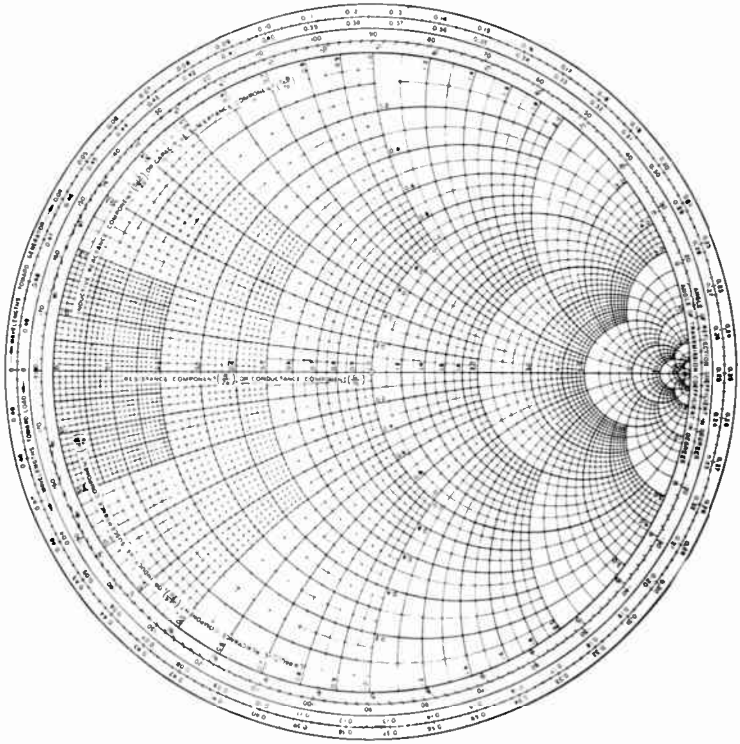


Fig. 4-1. The Smith chart, in full detail.

Several variations of the Smith chart have been developed for particular applications. The form used throughout this book is called the *normalized-impedance* Smith chart.

NORMALIZED IMPEDANCE

To use the same chart with transmission lines of different characteristic impedances, we use a *normalized impedance*. This is simply a value of impedance that has been divided by the characteristic impedance of the transmission line we are using. For example, suppose that we are working with a system where the transmission line has a characteristic impedance of 50 ohms—a very common value in broadcasting. To normalize the impedances in the system, we would simply divide them by 50 ohms.

Suppose that we have a load impedance that has been measured on a bridge and found to be 200 + j150 ohms. To normalize to 50 ohms, we would merely divide both the resistive and reactive parts by 50 ohms.

$$z = \frac{200}{50} + j\frac{150}{50} = 4 + j3$$

If the impedance was stated in polar form, we would simply divide the magnitude by 50 ohms, doing nothing to the phase angle. For example

$$Z = 200 + j150 = 250 \angle 36.87$$

$$z = 4 + j3 = \frac{250}{5} \frac{\angle 36.87}{3} = 5 \angle 36.87$$

If we had a load impedance that was a pure resistance of 50 ohms, when normalized it would be simply 1.

Note that whereas we use the capital letter Z to designate an impedance, we use a lowercase z to denote a normalized impedance. Actually, what we call a normalized "impedance" is not an impedance at all. It is simply a *ratio*, a pure number that doesn't have any units such as ohms. When we normalize an impedance—say, 100 ohms—we divide by so many ohms, as shown below.

$$\frac{100\text{ ohms}}{50\text{ ohms}} = \frac{100}{50} = 2$$

Note that the units *ohms* cancel out in the equation. The normalized value of impedance is simply 2, not 2 ohms.

This point might be a little confusing at first, but a rather silly example will make it clear. Suppose that we wish to compare the numbers of apples in some baskets. If the basket that we use as a reference contains 50 apples and another basket contains 100 apples, the ratio of the two is

$$\frac{100\text{ apples}}{50\text{ apples}} = 2$$

and not 2 *apples*. Similarly, a normalized impedance of 2 simply means that the impedance in question has twice as many ohms as the value to which we normalized it.

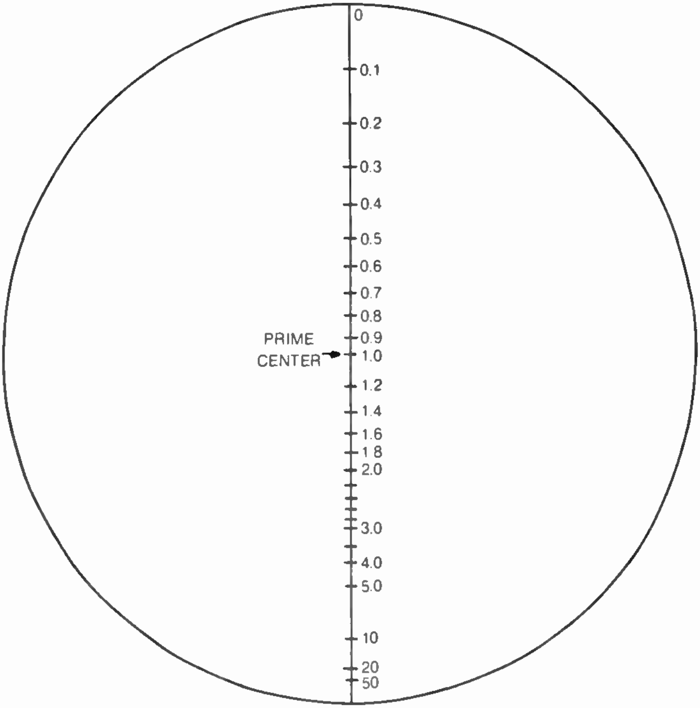


Fig. 4-2. Resistance axis of Smith chart.

Before we can use a normalized value of impedance to solve any circuit problems, we must reverse the normalizing procedure by *multiplying* by the normalizing value, (characteristic impedance), which in our example was 50 ohms.

The normalized notation is sometimes called a *per unit* notation. In the example 100-ohm impedance has a per-unit value of 2 when referred to 50 ohms. This means simply that there are 2 ohms for each ohm in the 50-ohm normalizing value.

When we normalize an impedance such as 200 + j150 ohms, we get a complex number that is proportional to the magnitude of the original value and has the same phase angle.

RESISTANCE SCALES

The first line of the Smith chart that we will consider is the resistance axis (Fig. 4-2). This is the only straight line on the entire chart. The center of this line, which is called the *prime center* of the chart, is labeled 1.0. It corresponds to a normalized value of resistance of unity (1). If we had a 50-ohm pure resistance and were using a 50-ohm transmission line, we would have a normalized value of 1 and would represent it by placing a dot at the prime center of the chart.

Below the prime center of the chart are points corresponding to normalized values of resistance greater than 1, with the bottom of the chart corresponding to infinity. Thus 100 ohms (normalized to 50 ohms) would be represented by a dot at the point labeled 2.0.

Above the prime center are normalized values less than 1. A 25-ohm resistance (normalized to 50 ohms) would be represented by a dot at the point labeled 0.5 on the resistance axis.

The resistance axis of the Smith chart is one axis of a graph, just as Y-axis of a rectangular graph is, but in the Smith chart the scales of the graph are circles rather than straight lines. The resistance scales are the circles shown in Fig. 4-3.

A value of normalized resistance is assigned to each circle. The largest circle, which coincides with the outer edge of the chart, corresponds to 0; and a dot at the bottom of the chart corresponds to infinity. Thus *any point* on the circle labeled 1.0 corresponds to a normalized-resistance value of 1.

REACTANCE SCALES

The reactance scales, which appear as curved lines in Fig. 4-4, are actually parts of circles. All of these lines are tangents to the resistance axis, which itself is the zero-reactance line. The circle that forms the outer edge of the chart can be thought of as the reactance axis of the chart.

Each reactance line is assigned a value of normalized reactance, which is labeled near the outer edge of the chart. Reactance lines to the right of the resistance axis are used for positive or inductive reactance, and those to the left of the resistance axis are used for negative or capacitive reactance. Thus an inductive reactance of 100 ohms (normalized to 50

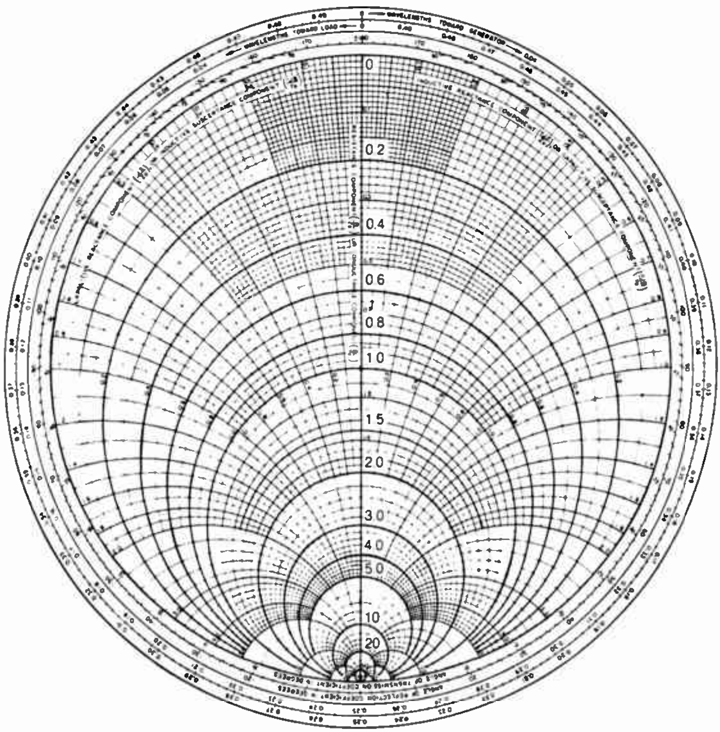


Fig. 4-3. Resistance scales.

ohms) would be represented by a dot on the reactance line labeled 2.0 on the right side of the chart.

PLOTTING IMPEDANCES

Now that we have a graph with both resistance and reactance scales, we can plot various values of impedance as points on the graph. Since the scales of our graph are in normalized values of impedance, we must normalize each impedance before we plot it. For the remainder of this chapter we will assume that we are working with a transmission line having a characteristic impedance of 50 ohms, and we will normalize all impedances to this value.

Figure 4-5 shows several impedances plotted on the coordinates of a Smith chart. The absolute values, normalized values, and locations of these impedances are tabulated as:

Absolute Impedance, Ohms	*Normalized Value*	*Point on Fig. 4-5.*
50	1	A
50 + *j*100	1 + *j*2	B
50 − *j*100	1 − *j*2	C
100 + *j*100	2 + *j*2	D
100 − *j*100	2 − *j*2	E

Two points on the Smith chart are of particular interest in connection with solving certain transmission-line problems. The first is the impedance of an ideal short circuit. Here, both the resistance and the reactance are zero. The value of impedance is represented by a dot at the top of the chart, where the resistance and reactance axes intercept (point *F* in Fig. 4-5). The other point of particular interest is the impedance of an ideal open circuit. Here the resistance is infinite and the reactance is zero. The resistance portion of the

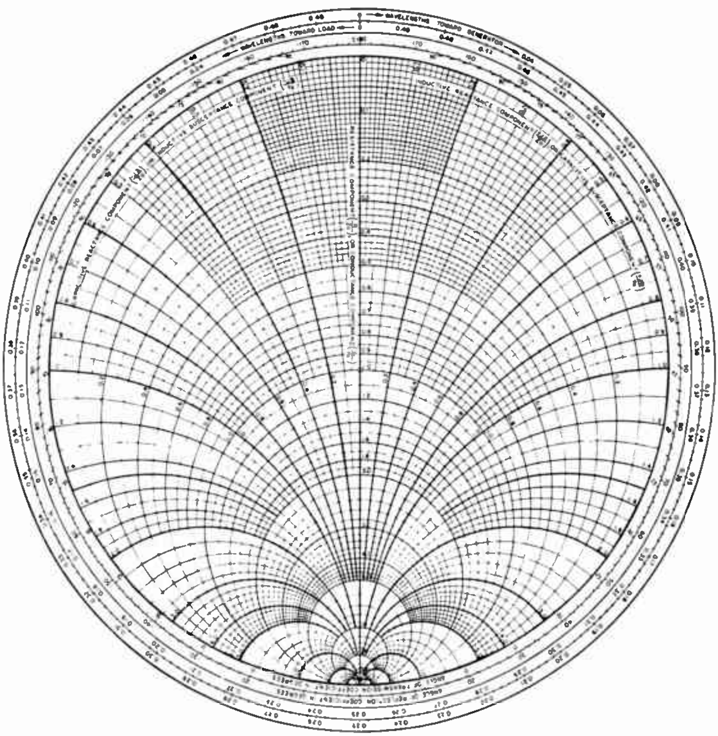


Fig. 4-4. Reactance scales.

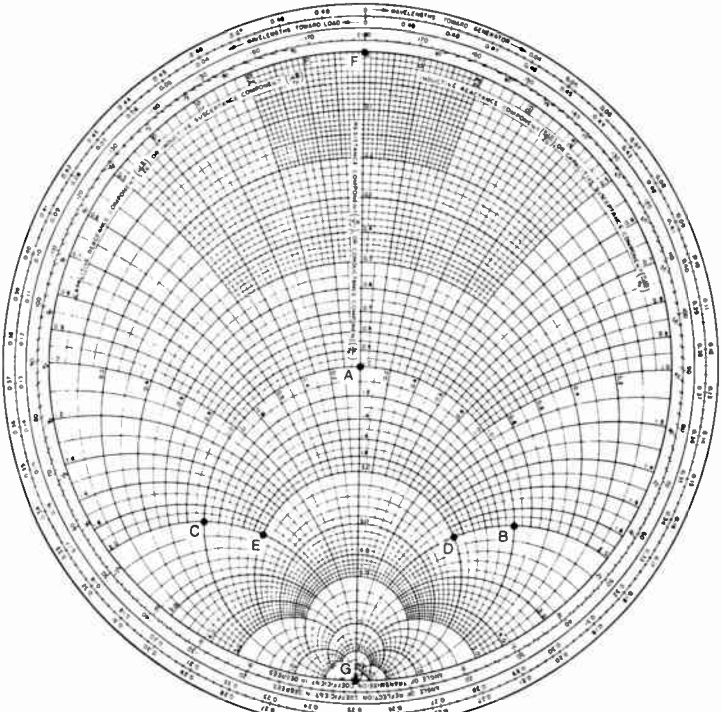


Fig. 4-5. Impedances plotted from table in text.

normalized value is infinity, because an infinitely large number divided by 50 is still infinitely large. Thus the impedance of an ideal open circuit is represented by a dot at the bottom of the chart (point *G* in Fig. 4-5).

To summarize, any value of normalized impedance can be plotted on the Smith chart. Pure resistances fall on the resistance axis, the vertical line through the center of the chart. Pure reactances fall on the zero-resistance circle, around the outer edge of the chart. Complex impedances having both resistance and reactance fall somewhere on the face of the chart. Impedances having an inductive reactance lie on the right half of the chart, and those having capacitive reactance lie on the left half of the chart.

VSWR AND WAVELENGTH SCALES

So far we have shown that any value of impedance can be plotted on a Smith chart. We haven't however, justified the use of circular rather than rectangular scales. We will do this now.

We saw in Chapter 3 that in a lossless transmission line not terminated in its characteristic impedance, impedance measured along the line varies with distance from the load. We also saw that the impedance repeats itself every 1/2 wavelength along the line. Now we can see one of the advantages of the choice of coordinates on the Smith chart. All of the values of impedance measured along a transmission line will fall on a circle—the VSWR circle—on the chart.

Suppose we have a 50-ohm transmission line that is terminated in a 100-ohm resistance. Using a formula from Chapter 2 we find that the standing-wave ratio is

$$\text{VSWR} = \frac{R_L}{Z_0} = \frac{100}{50} = 2$$

Now, if we draw a circle centered at the prime center of the chart, with a radius equal to 2 on the resistance axis (see Fig. 4-6), all values of impedance that can be measured along the

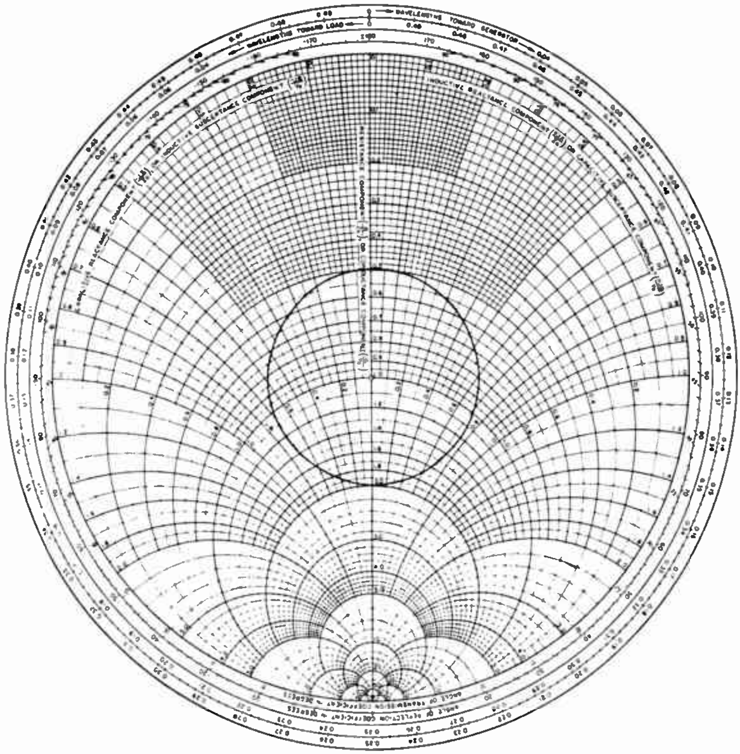


Fig. 4-6. A VSWR circle for VSWR = 2.

line will fall on this circle. This will hold true as long as the losses along the line are negligible.

The standing-wave circles are usually not printed on the chart. They are drawn for individual cases by the user of the chart. Remember that the radius of the circle is equal to the numerical value of the VSWR measured on the resistance axis of the chart.

Looking at the VSWR circle in Fig. 4-6, we see that once we go around the circle from any point, we are right back to the impedance with which we started. This corresponds to moving along a transmission line a distance of 1/2 wavelength. Thus we can conclude that going around the Smith chart once is comparable to moving along a transmission line a distance of 1/2 wavelength.

Wavelength scales are provided along the outer edge of the chart and are marked in decimal fractions of a wavelength. The outer scale, which increases in a clockwise direction, represents distance along the line in the direction of the generator, that is, away from the load. The inner wavelength scale is marked in decimal fractions of a wavelength toward the load.

With the scales that we have described, we can find the impedance at any point along a line, as well as the standing-wave ratio, if we know the load impedance and the characteristic impedance of the line. Suppose, for example, that we have an impedance of $50 + j50$ ohms and a 50-ohm transmission line. The normalized value of load impedance is thus $1 + j1$, which is represented as a dot at point *A* of Fig. 4-7. We can then draw a VSWR circle centered about the prime center of the chart and passing through point *A*.

We can read the standing-wave ratio directly from the lower point on the resistance axis (point *B*), where the VSWR circle intercepts it. In this case, the VSWR is about 2.6. This point also corresponds to the impedance at a point about 0.09 wavelength from the load, as shown on the wavelength scale of the chart.

A careful inspection of the VSWR circle in Fig. 4-7 shows many interesting things. At points *B* and *C*, where the circle crosses the resistance axis, the reactance is zero. At point *C* the impedance is not only purely resistive, but it has its lowest value. This point, therefore, must correspond to the point on the line where the voltage of the standing wave is minimum and the current is maximum. Point *B*, which is 1/4 wavelength from *C*, corresponds to the point along the line where the voltage is maximum and the current is minimum.

So far we have only dealt with conditions over a single half-wavelength of line from the load. Since we are considering the line to be lossless, we can find the impedance at any point along the line by merely going around the VSWR circle once for each half-wavelength of distance from the load.

In practice, this is accomplished by merely subtracting the largest possible number of half-wavelengths from the line length. For example, if we have the situation of Fig. 4-7, where the normalized load impedance is $1 + j1$, and we wish to find

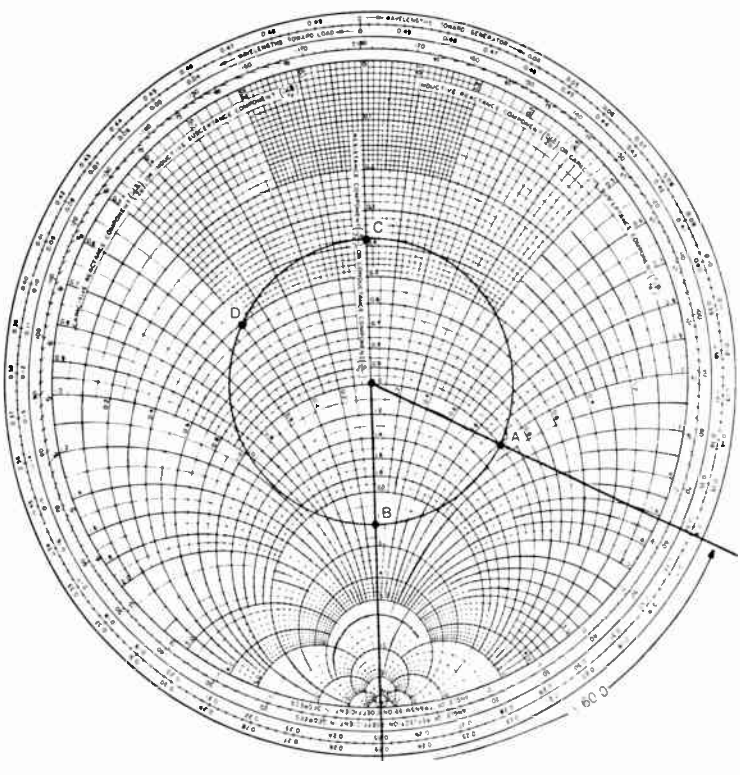


Fig. 4-7. All values of impedance along a line, however long, can be found on the Smith chart.

the impedance at a point 1 3/4 wavelengths away, we merely subtract 1 1/2 wavelengths from the total and find the impedance 1/4 wavelength from the load. This is at point *D* in Fig. 4-7. The impedance at point *D* is $0.5 - j0.5$. This is the reciprocal of the normalized load impedance, which is just what we would expect, since a quarter-wave line inverts the load impedance.

Earlier we saw that the input impedance of a lossless quarter-wave line is given by

$$Z_{in} = \frac{Z_0^{\ 2}}{Z_L}$$

This equation is based on using absolute values of impedance in ohms. We can rearrange it to use normalized values of impedance by dividing both sides by Z_0, giving us

$$\frac{Z_{in}}{Z_0} = \frac{Z_0^{\ 2}}{Z_L\ Z_0} = \frac{Z_0}{Z_L}$$

If we use a lowercase *z* to represent normalized values, we have simply

$$z_{in} = 1/z_l$$

and the characteristic impedance of the line cancels out of the equation completely. This is another convenience that results from using normalized or per-unit values.

RADIALLY SCALED PARAMETERS

Depending on the type of Smith chart, there are as many as eight different radial scales (Fig. 4-8). Each of these scales starts at the prime center of the chart and extends radially outward. In the Smith chart calculator these scales are printed on a plastic cursor that is pivoted at the prime center of the chart. In the printed charts the radial scales are printed along one side or along the bottom of the chart (Fig. 4-9).

The SWR scale gives the voltage or current standing-wave ratio (Fig. 4-9). The scale may be used on a lossless transmission line by drawing a line from the standing-wave circle to the scale, as shown in Fig. 4-9, or by using a compass to transfer the distance between the prime center and standing-wave circle to this scale. The SWR scale is really not

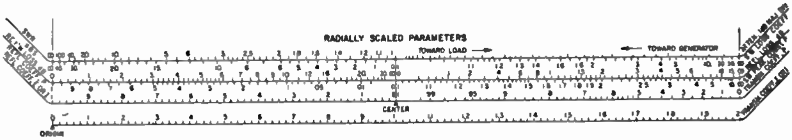


Fig. 4-8. The radial scales of the Smith chart are located along one edge. SWR scale is useful for determining SWR on lossy lines. The reflection coefficient scale, together with the angle of reflection coefficient scale on the circular chart, is useful for finding the magnitude of the reflected voltage as a function of the forward voltage, also the angle between the forward and reflected voltages.

necessary, particularly on lossless lines, because the standing-wave ratio can be read directly from the resistance axis below the prime center. The scale is useful in computations for lossy lines because, as we shall see later, the SWR curve on the chart is not a circle for such lines.

Opposite the SWR scale, the VSWR is expressed in decibels. This parameter is sometimes useful. It is the ratio in decibels of the maximum to minimum voltage or current along

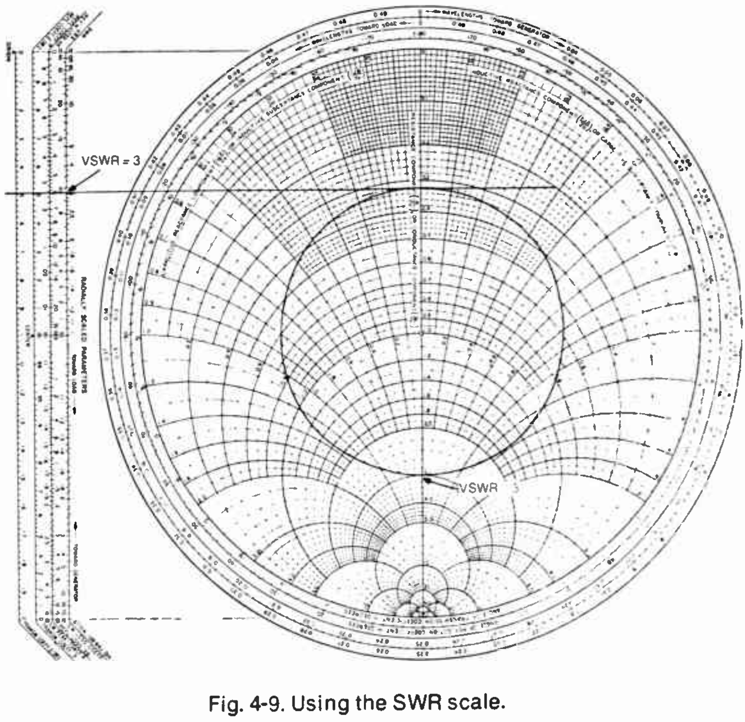


Fig. 4-9. Using the SWR scale.

a transmission line and is given by

$$\text{VSWR, dB} = 20 \log \frac{E_{max}}{E_{min}}$$

Another radial scale, together with a scale on the chart which we haven't considered yet, can be used to find both the magnitude and the angle of the reflection coefficient. Figure 4-10 is for the same load as Fig. 4-7, namely, $1 + j1$. The radial line drawn from the prime center of the chart, through the load impedance to the peripheral scale labeled *angle of reflection coefficient*, shows the angle to be about 63.4°. We can find the magnitude of the reflection coefficient by measuring the length of a line from the prime center of the chart to the load impedance ($1 + j1$) and transferring the distance to the radial scale marked *reflection coefficient*. The reflection coefficient in this case is about 0.45. This information gives us the magnitude of the reflected voltage as a function of the forward voltage, and the angle between the forward and reflected voltages.

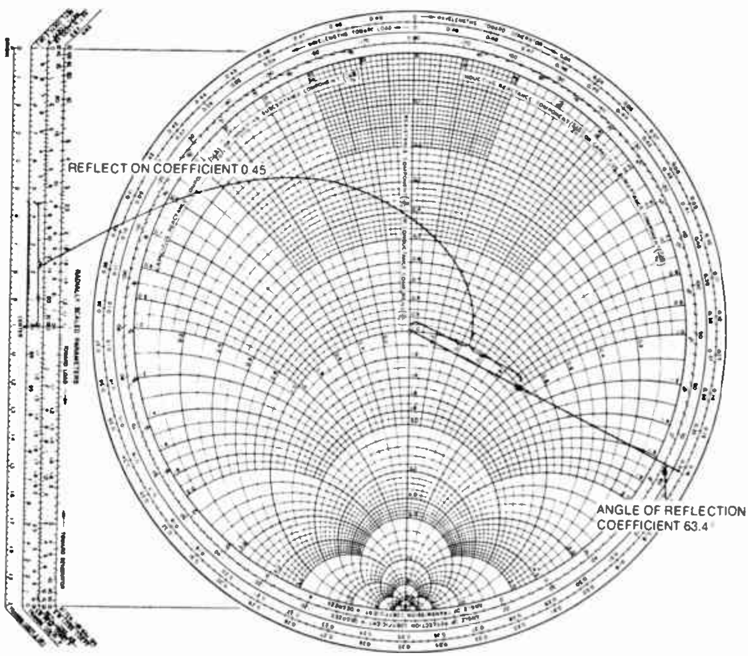


Fig. 4-10. Finding the reflection coefficient.

Another radial scale on most charts is for the *power reflection coefficient*. We shall have no occasion to use this scale.

LINE LOSSES

So far we have considered that transmission-line losses are low enough that we can ignore them. In a great deal of broadcast work, this is entirely practicable. There are cases, however, where line losses are significant, and radial scales of the Smith chart provide a way of handling them without many tedious mathematical considerations.

One effect of transmission-line loss is that the standing-wave ratio is not constant along the line. In a lossy line the reflected wave is attentuated as it travels back toward the sending end, so it will naturally have less effect on the voltage distribution along the line. Thus the standing-wave

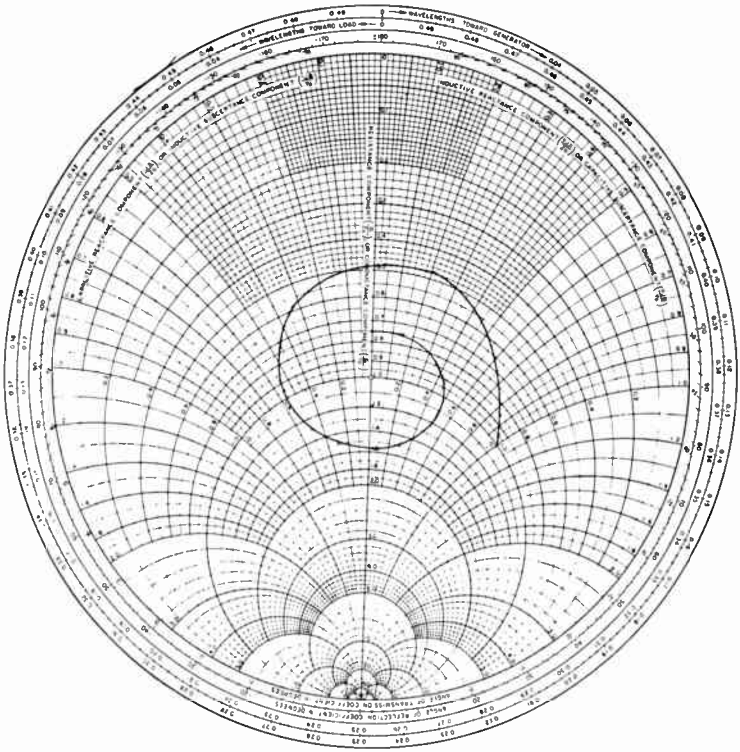


Fig. 4-11. The SWR curve for a lossy line.

ratio becomes smaller in a lossy line as we move closer to the source. For a lossy line the standing-wave curve is no longer a circle; it is a curve that starts at the load impedance and spirals clockwise toward the prime center of the chart, as shown in Fig. 4-11.

How close the curve comes to the center depends on the amount of loss in the line. In an extreme case the spiral would come very close to the prime center of the chart. This would mean that the loss of the line was so great that the reflected wave would have no effect on the input impedance, which would then be equal to the characteristic impedance of the line.

In practice, it is rather difficult to construct the spiral for a lossy line on a Smith chart. Fortunately it isn't necessary. There is a radial scale, *transmission loss, 1 dB steps*, that can be used. This scale (Fig. 4-12) is a *relative* scale, so the divisions are not numbered. It is sufficient to know that the divisions are 1 dB apart.

Assume, for example, that we measure a normalized impedance of $1.5 + j2$ at the terminals of a transmission line of 3/8 wavelength and we know that the loss of the line is 1 dB. We want to know the VSWR at both the input and load, as well as the load impedance.

We start out just as we would with a lossless line, by drawing a VSWR circle through the measured impedance. This is the inner circle in Fig. 4-12, drawn through point *A*. We read the VSWR at this end of the line from point *A* on the figure and find it to be about 4.6. Now, to find the VSWR at the load, we use the radial 1 *dB step* scale.

Transferring the radius of our VSWR circle to the *1 dB step* scale, we find that the intercept is at about the second major division from the outer edge of the scale, as shown in the figure. Since the loss in the line is 1 dB, we move up to the next major division to form a new radius as shown. With this new radius, we can draw a new VSWR circle. This new circle, which passes through the load impedance, shows that the VSWR at the load is about 9. Inasmuch as the length of the line is 3/8 wavelength, we can find the load impedance by moving 0.375 wavelength around the new circle toward the load. We

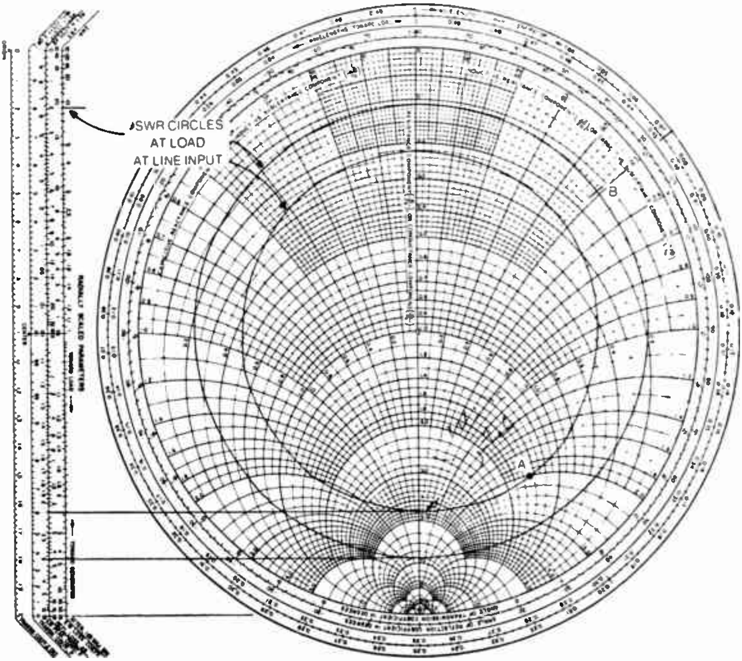


Fig. 4-12. SWR circles for a lossy line.

thus find that the actual load impedance is about 0.1 + *j*0.5 (point *B*).

This technique of drawing two circles to solve a problem is much easier than drawing a spiral, but remember: Each circle is valid at only one point along the line, whereas with a lossless line, one circle applies all along the line.

An inspection of the *1 dB steps* scale will provide a great deal of insight into how losses affect the standing-wave ratio on a line. If we start at the end of the scale, corresponding to the prime center of the chart, we see that the 1 dB steps are spaced very closely. Thus, if the VSWR at the load is small, even several decibels of loss in the line will not have an extremely great effect on the VSWR. If the VSWR at the load was 1.4 and the loss of the line was 3 dB, the VSWR at the input to the line would be 1.2.

On the other hand, at large distances from the prime center of the chart, the decibel steps are spaced much farther apart. Thus, with large values of VSWR, even a small amount of line loss will have a large effect. If the VSWR at the load was 10 and the line loss was 1 dB, the VSWR at the input would be less than 5.

We can see how this happens. When the VSWR at a load is low, there isn't much energy reflected, so the losses can't have much effect on it. If the VSWR is large, a lot of energy is reflected, so the losses will have more effect on it.

SMITH CHART PLOTS FOR VARIOUS TERMINATIONS

When a line is terminated in its characteristic impedance, the normalized impedance of the load is 1, so the standing-wave ratio is 1, and the behavior of a lossless line so terminated is represented by a dot at the prime center of the chart, point *A* in Fig. 4-13.

If a line is open at the receiving end, the load impedance is represented by a dot at the bottom of the chart, point *B* in Fig.

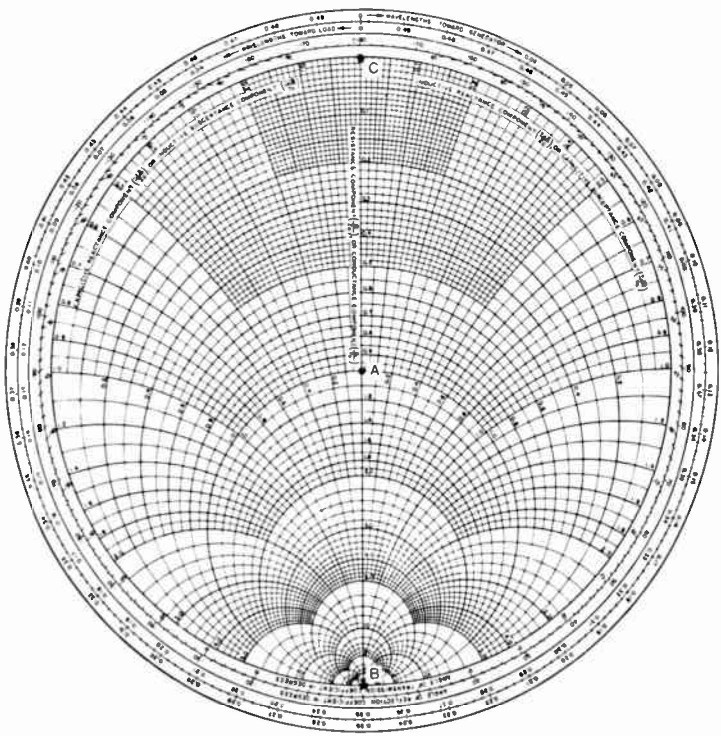


Fig. 4-13. Smith chart plots for various terminations.

4-13. The standing-wave ratio is infinite and is represented by a circle around the outer edge. The impedance at any point along this line is found by moving clockwise the appropriate fraction of a wavelength around the circle. Note that the circle coincides with the zero-resistance circle, which means that as long as there are no losses in the line (which would cause our circle to become a spiral), the input impedance will be a pure reactance. As we move up along this circle, we find that the input impedance is a high capacitive reactance. This decreases until—when we reach the top of the chart, which corresponds to a distance of 1/4 wavelength from the open-circuited end—we find that the input impedance is a short circuit. This of course is what we would expect, since we know that an open quarter-wave line looks like a short circuit.

As we continue around the chart, we see that the input impedance becomes an inductive reactance, first small in value, but becoming larger until, when we get to a point 1/2 wavelength along the line, we are right back where we started. The input impedance is then an open circuit.

If the receiving end of the line was terminated in a short circuit, we would have the same VSWR circle, corresponding to an infinite standing-wave ratio, but we would start at point *C* of Fig. 4-13 and again move clockwise around the chart to find the impedance at various points along the line. This demonstrates that a shorted line behaves the same as an open line except that the same conditions are displaced by a 1/4 wavelength along the line.

These examples show how familiarity with a Smith chart reinforces a knowledge of transmission-line principles and simplifies computations.

GETTING IN AND OUT OF THE CHART

We have looked over each of the scales on the Smith chart and seen how it relates the load impedance, impedance along a line, reflection coefficient and standing-wave ratio. Nevertheless, there are still a lot of lines on the chart, and until one has had quite a bit of practice, the chart is apt to be confusing. It is helpful to use some systematic method of plotting points on the chart and reading the desired results.

The way we enter a number on the Smith chart depends, of course, on the type of quantity. An impedance isn't entered on the chart in the same way as the angle of a reflection coefficient, for example. When we use the Smith chart, we know some quantities and use the chart to find others. The quantities that we start with come from either basic design considerations or from measurements that we can make.

There are several different types of instruments for measuring transmission-line parameters. The particular measurement that is made depends on the availability of instruments and the application. For example, in the standard broadcast band, we frequently measure impedances on a bridge, whereas, in an FM or TV station, we might measure the reflection on a line. The beauty of the Smith chart is that it relates all of the transmission-line parameters.

One parameter that fortunately is nearly always known in advance is the characteristic impedance of the transmission line. This means that normalizing any particular value of impedance is simply a matter of division.

Frequently the quantity we start with is the value of an impedance. Using lowercase letters to represent normalized values of resistance, reactance, and impedance, we can express our number in the form $z = r + jx$. Suppose, for example, that it happens to be $1.4 + j2$. To enter this value on the Smith chart, we start with the resistance. Since the top of the resistance axis appears at the top of the chart, we start there. Starting at point A in Fig. 4-14, we move down the resistance axis until we find the circle corresponding to 1.4. This is at point B. We know that the value we want to plot lies somewhere on this circle. To find out which way to go along this circle, we look at the sign of the reactance in the impedance we are plotting. In our example the sign is positive. This means that the impedance lies on the right half of the chart, so we move to the right along the 1.4 resistance circle until we come to the 2.0 reactance circle, point C in the figure. This procedure probably isn't necessary for one who is thoroughly familiar with the chart, but it is convenient and reduces the probability of error.

If the impedance that we have just plotted is either a load impedance, or the impedance measured at some point along a

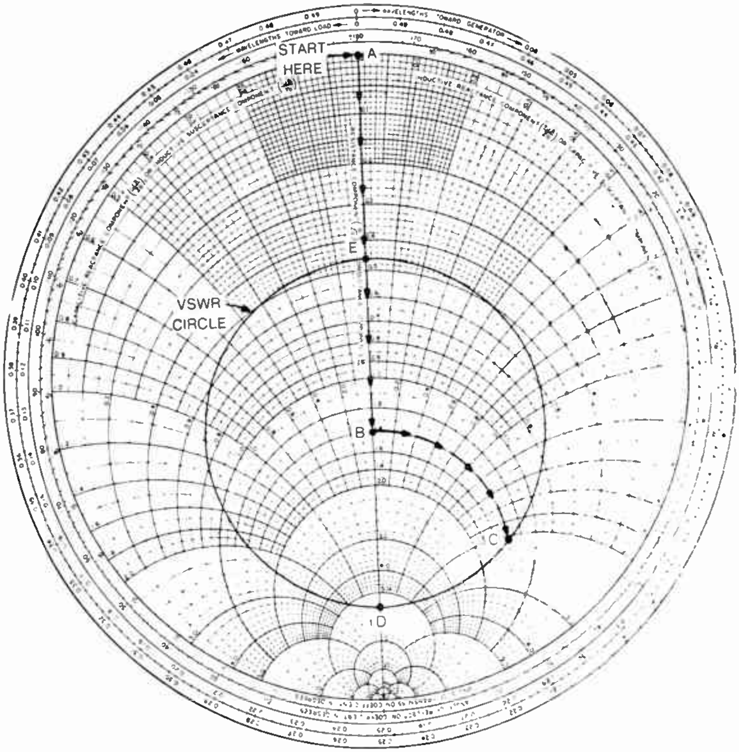


Fig. 4-14. Plotting data on the Smith chart.

line that we can consider lossless, we can then determine many of the other parameters. We can construct a standing-wave circle and read off the VSWR. If it is a load impedance, we can also read off the magnitude and angle of the reflection coefficient.

If the plotted impedance is the value measured at some point along the line, we need one more piece of information before we can find the value of the load impedance: either the distance in wavelengths from the point of measurement to the load, or the angle of the reflection coefficient. If we know that the load is a pure resistance but don't know what the value is, we can narrow it down to two values. A pure resistance would have to fall on the resistance axis, so its value is that at either point *D* or point *E* in Fig. 4-14.

If one of the quantities that we know to start with is the standing-wave ratio, we can start by plotting the VSWR circle.

This then tells us all possible values of impedance along the line, but it doesn't tell us where they occur. We know that both the load impedance and input impedance are on this circle, but we must have more information before we can find anything more specific. In laboratory-type measurements with a slotted line, we can often find the distance in wavelength from the load to the first maximum or minimum of the standing wave. Since we know that the maximum and minimum of the standing wave occur where the VSWR circle crosses the resistance axis, we can use the wavelength scales to find the value of the load impedance.

ADMITTANCE PARAMETERS

Admittance, conductance, and susceptance are much less familiar concepts than impedance, resistance, and reactance, but no more difficult inherently. Admittance and related parameters were explained in Chapter 1. Admittance is reintroduced at this point for two reasons:

1. There are situations in antenna and transmission-line work where using admittance rather than impedance considerably simplifies computations.
2. The Smith chart greatly simplifies conversion from impedance to admittance, and vice versa.

ADMITTANCE AND THE SMITH CHART

If the admittance concept seems fearsome, the Smith chart will come to the rescue. First let us use the chart to find a value of admittance corresponding to a given value of impedance. To keep things simple, let's find the admittance corresponding to a resistance of 100 ohms. Normalizing this to 50 ohms gives us a value of 2.0, which we represent by a dot at point *A* of Fig. 4-15.

Now, we know that the normalized impedance seen looking into a quarter-wave transmission line is equal to the reciprocal of the normalized load impedance. We also know that this reciprocal is equal to the normalized load admittance. So all we have to do is go around the chart 180°, which corresponds to moving along a transmission line 1/4 wavelength, and read 0.5

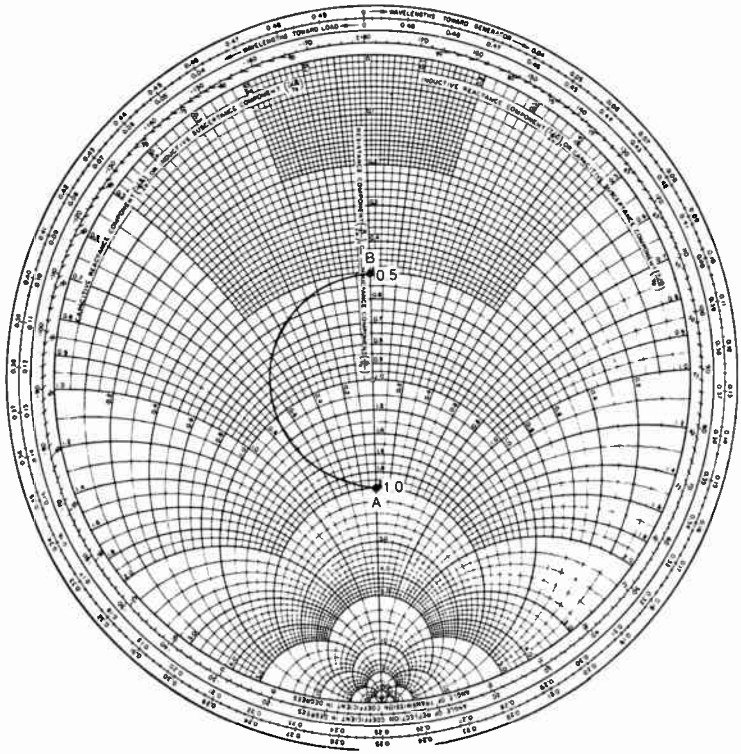


Fig. 4-15. Finding admittance on a Smith chart.

at point *B*. The value 0.5 represents the *normalized* admittance of the load. To get back to an *absolute* value of admittance in mhos, we merely have to divide by 50 ohms, or what is the same thing, multiply by 0.02 mho.

$$\lambda = 0.02y = 0.02 \times 0.5 = 0.01 \text{ mho}$$

Of course, in this simple example it would have been easier to divide 100 ohms into 1, but when the impedance is complex, the Smith-chart method is easier. For example, to find the admittance corresponding to an impedance of $90 - j70$ ohms by the usual method, we would have to rationalize the expression and go through the following steps:

$$Y = \frac{1}{Z} = \frac{1}{90 - j70} \times \frac{90 + j70}{90 + j70} = \frac{90 + j70}{8100 + 4900}$$

$$= \frac{90}{13000} + j\frac{70}{13000} = 0.007 + j0.005 \text{ mho}$$

To find the admittance corresponding to an impedance of $90 - j70$ ohms using the Smith chart, we must first normalize to some reference value of impedance. To keep the math simple, let's use 100 ohms. The normalized value of impedance is then $0.9 - j0.7$. This value is entered at point *A* of the Smith chart of Fig. 4-16, and the normalized value of admittance is read off the chart at point B, 180° away, at an equal distance from the prime center of the chart. The normalized value of admittance is $0.7 + j0.5$. To obtain the unnormalized admittance, we divide by 100 (or multiply by 0.01), which gives us $0.007 + j0.005$ mhos—the same value that we obtained earlier with considerable calculation. The positive sign of the susceptance means it is capacitive.

EXPANDED-SCALE CHARTS

One of the advantages of the Smith chart over other impedance graphs is that all values of normalized impedance

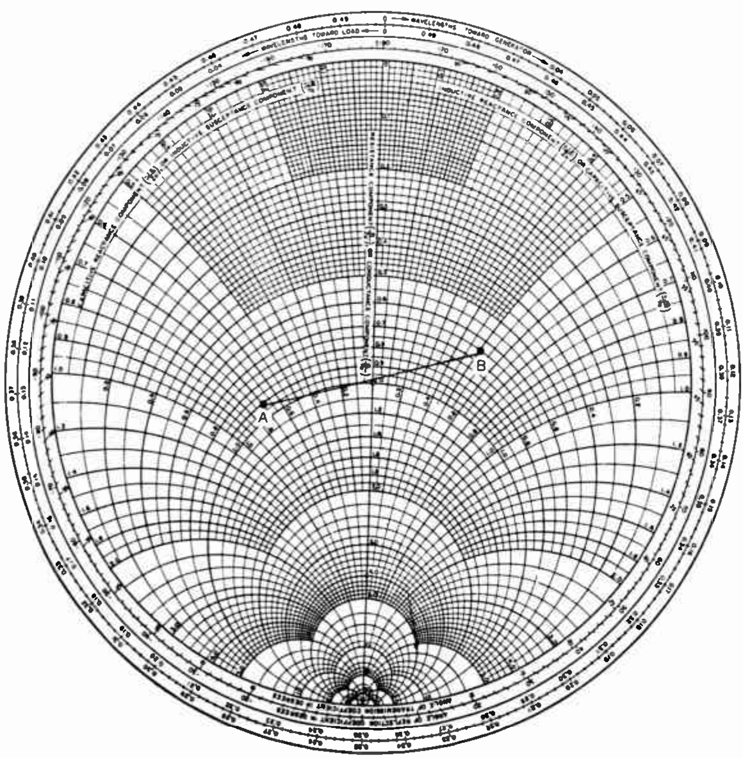


Fig. 4-16. Example of admittance calculation.

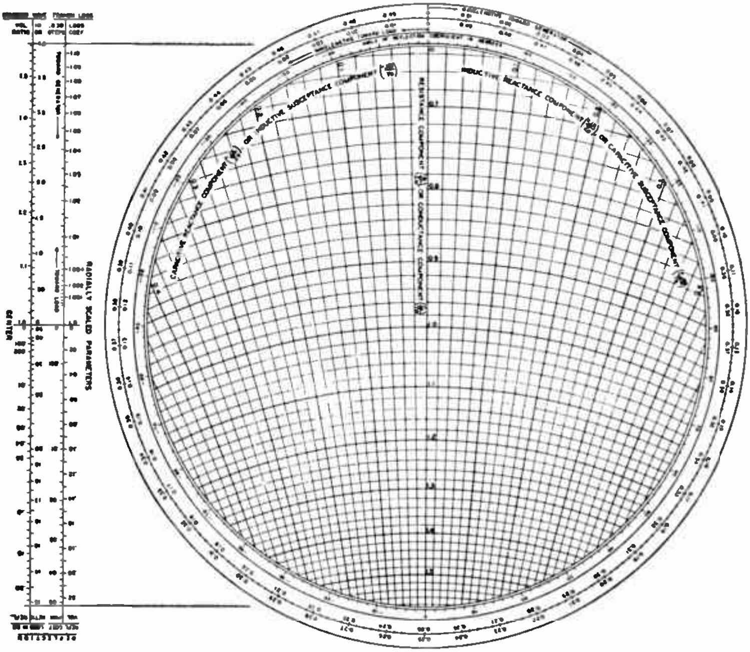


Fig. 4-17. Expanded-scale Smith chart.

from zero to infinity can be plotted on the Smith chart. One of the penalties we pay for this convenience is an inherent inaccuracy in reading the chart. This isn't very significant when the mismatch and hence the standing-wave ratio is high. It *is* annoying in a closely matched system, where the standing-wave ratio is held to close limits. For such cases, expanded charts are available, as shown in Fig. 4-17. Here the area near the prime center of the regular chart is expanded to cover the entire area of the regular chart. The scales are expanded radially about the prime center by a ratio of 4.42:1. On this expanded chart the magnitude of the reflection coefficients from 0 to 1/4.42 can be displayed, whereas, on the full chart, all possible values of reflection coefficient from 0 to 1 can be plotted. The maximum standing-wave ratio that can be plotted without going off the edge of the expanded chart is 1.59, which, in decibel notation, is 4 dB.

The expanded chart is excellent for plotting transmission line and load-impedance characteristics in systems where the standing-wave ratio does not exceed 1.59.

Smith charts with other degrees of expansion are available. In fact, a wide variety of special charts have been made at different times for particular applications. The two that have been described here are the ones most commonly used for broadcast work.

USING THE SMITH CHART

Although the main body of the Smith chart has only two sets of scales, for resistance and reactance, it can be used for many different purposes.

Analyzing Networks

The Smith chart can be used for analyzing networks because it provides an easy way of converting between impedances and admittances. Combining admittances in parallel is simply a matter of addition, the same as combining impedances in series. Figure 4-18A shows an *L*-network of the type used to match impedances in antenna systems. To analyze this network on a Smith chart, we must first normalize all of the impedances, preferably to the value of the characteristic impedance of the transmission line—in this case, 50 ohms. The normalized impedances are shown in Fig. 4-18B. We can now add the impedances and admittances of the network directly on a Smith chart.

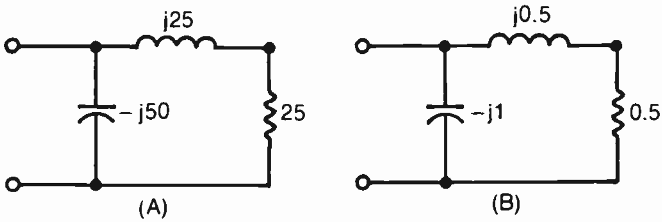


Fig. 4-18. An L-Network, with actual and normalized impedances.

The impedance of the load is a normalized resistance of 0.5. Therefore we enter the Smith chart from the top and go to point *A* in Fig. 4-19A, which represents a normalized impedance of 0.5. Next we add the series reactance of $+j0.5$. This means moving along the 0.5 resistance circle to point *B*. The next element that we wish to consider is the capacitor. Inasmuch as the capacitor is connected in parallel with the rest of the network, we convert the reactance of the capacitor to a susceptance, the impedance of the network to an admittance, so that we can merely add them together.

The normalized impedance seen looking into the resistor and inductor in series is represented by point *B* in Fig. 4-19. We convert this normalized impedance to a normalized admittance by merely rotating the vector that extends from the prime center of the chart to point *B* through an angle of 180°. Or we can simply move through the prime center of the chart to a point that is diagonally opposite point *B*. Either way, we arrive at point *C*. Until now, we have been considering the chart of Fig. 4-19 as an impedance chart, where all of the numbers represent normalized resistances or reactances. We will now use the same chart as an *admittance* chart. All of the scales will represent normalized conductances or susceptances.

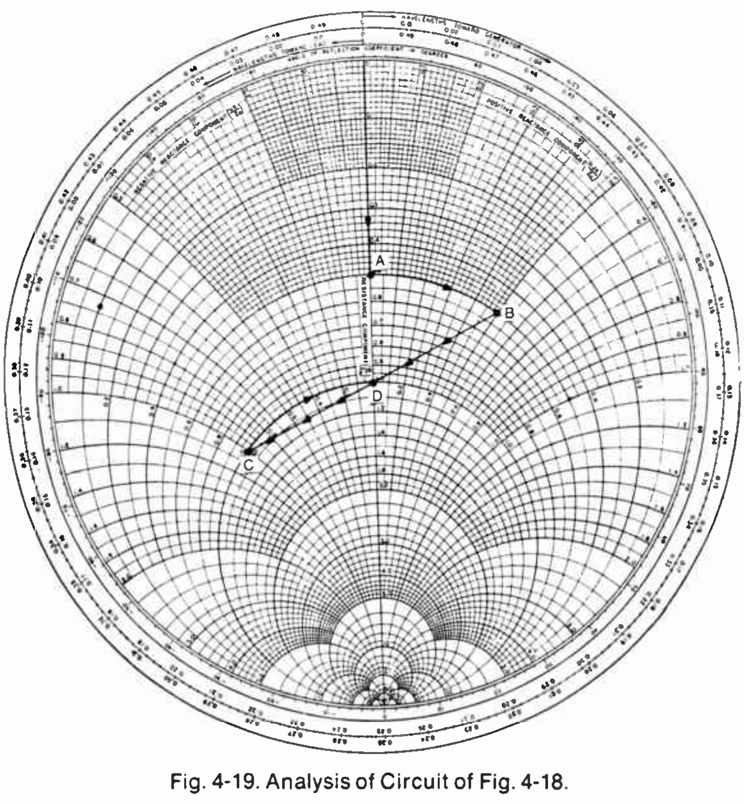


Fig. 4-19. Analysis of Circuit of Fig. 4-18.

Reading the chart as an admittance chart, we see that the admittance at point *C*, which is the admittance of the resistor and inductor in series, is $1 - j1$. We want to add the susceptance of the capacitor to this value of admittance. The susceptance of the capacitor is the reciprocal of its reactance. Since the normalized reactance in this example is $-j1$, the normalized susceptance will be $+j1$. We can add this susceptance by moving clockwise along the 1.0 conductance scale to point *D*. Here we see that the normalized admittance is simply 1.0. Converting from normalized values of admittance to absolute values, we must multiply by 0.02 mho, the reciprocal of 50 ohms. Thus the impedance seen looking into the network of Fig. 4-18 is 50 ohms.

At first the above procedure looks like a lot of work to get a simple answer. It is a rather lengthy procedure until one becomes familiar with it; then it is not only easy, but it gives increasing insight into network behavior.

Figure 4-20 shows another network that we may need to analyze. In this case, the resistor is connected in parallel with a capacitor, so we use conductance and susceptance values of these two elements, then convert to impedance to handle the series inductor. In Fig. 4-20B the resistance and reactances are normalized to 50 ohms. In Fig. 4-20C the resistive and

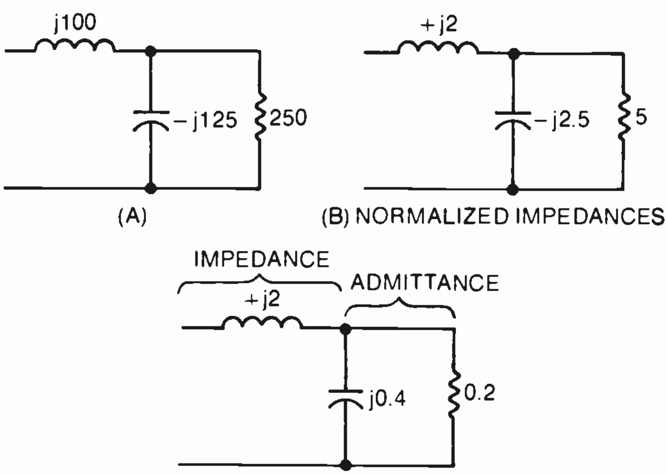


Fig. 4-20. The L-network referred to in the test.

capacitive values are given as a conductance and a susceptance because they are in parallel.

We start out by considering the Smith chart of Fig. 4-21 as an admittance chart. At point *A* we have the normalized conductance of the resistor, which is 0.2. Adding the capacitive susceptance of $+j0.4$ to this moves us to point *B*. The next element that we wish to consider is the inductor; it is connected in series, so we change the admittance to impedance by moving diagonally across the chart to point *C*. Here we see that the impedance of the resistor and capacitor in parallel is $1 + j2$.

From now on we consider the chart of Fig. 4-20 to be an impedance chart. Adding the inductive reactance, which is $+j2$ ohms, brings us to point *D*, which is at the prime center. Multiplying by our normalizing resistance of 50 ohms shows

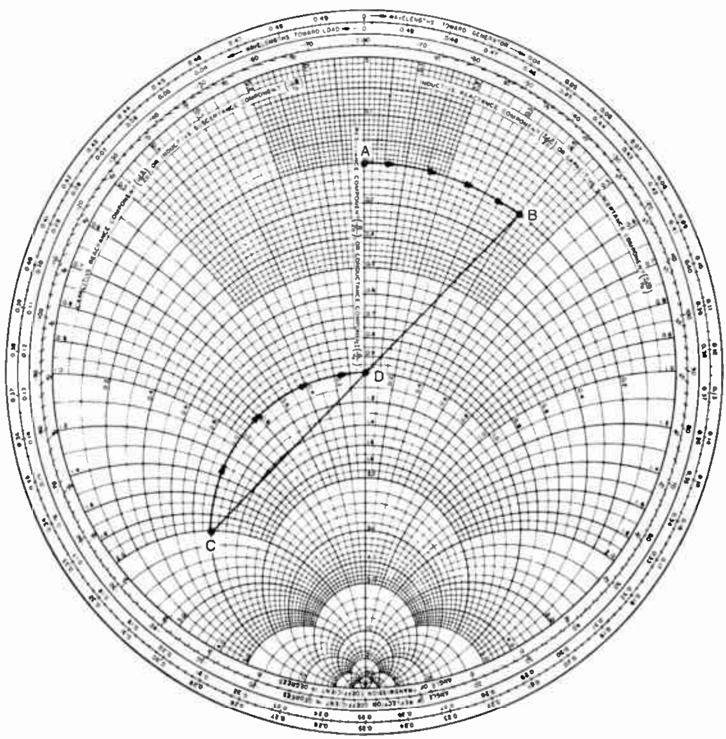


Fig. 4-21. Analysis of circuit of Fig. 4-20.

that the impedance seen looking into the network of Fig. 4-19A is 50 ohms.

Analyzing Data

Another very useful application of the Smith chart is in analyzing data from measurements. For example, suppose we have a transmission line that is 1/4 wavelength long at some frequency and is terminated in a given impedance. If we vary the frequency, the line will no longer be 1/4 wavelength long, and the driving-point impedance of the line will vary. We can use the Smith chart to show not only the driving-point impedance of the line but also how the standing-wave ratio varies with frequency.

Suppose, for example, that the 50-ohm transmission line shown in Fig. 4-22 is a quarter-wave line and is terminated by a 100-ohm resistor. Assume that we want to know the input impedance looking into the line at the frequency f_0, at which the line is a 1/4 wavelength long, and at two other frequencies, f_l and f_h, which are 10% lower and higher in frequency, respectively. First we normalize the 100 ohms to 50 ohms, giving a normalized resistance of 2.0.

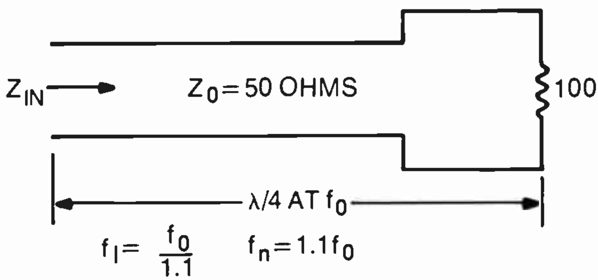


Fig. 4-22. Transmission-line problem.

We enter the Smith chart of Fig. 4-23 at this value, point *A*. To find the impedance at the sending end of the quarter-wave section, we merely move at a constant distance from the prime center of the chart, through 180°. Remember that 180° on the chart corresponds to 90° on the line itself. This brings us to point *B*, where we see that the normalized input impedance of the line is 0.5. At the frequency f_l, which is 10% lower than f_0, the line length will be 10% shorter than 1/4 wavelength. This is the same as moving 10% of a 1/4 wavelength, or 0.025 wavelength toward the load. This brings us to point C on the chart, which gives the normalized input impedance as $0.51 - j0.11$. Similarly, at the frequency f_h, which is 10% higher than f_0, the line will be 10% longer than 1/4 wavelength. This is at point D, where the normalized input impedance is $0.51 + j0.11$.

Multiplying all of the resistances and reactances by the 50-ohm normalizing resistance gives us the following input impedances at the three frequencies:

Frequency	Z_{in}
f_l	$25 - j5.5$
f_0	25
f_h	$25 + j5.5$

This shows that the resistive portion of the input impedance does not change very much as the frequency is changed by 10%

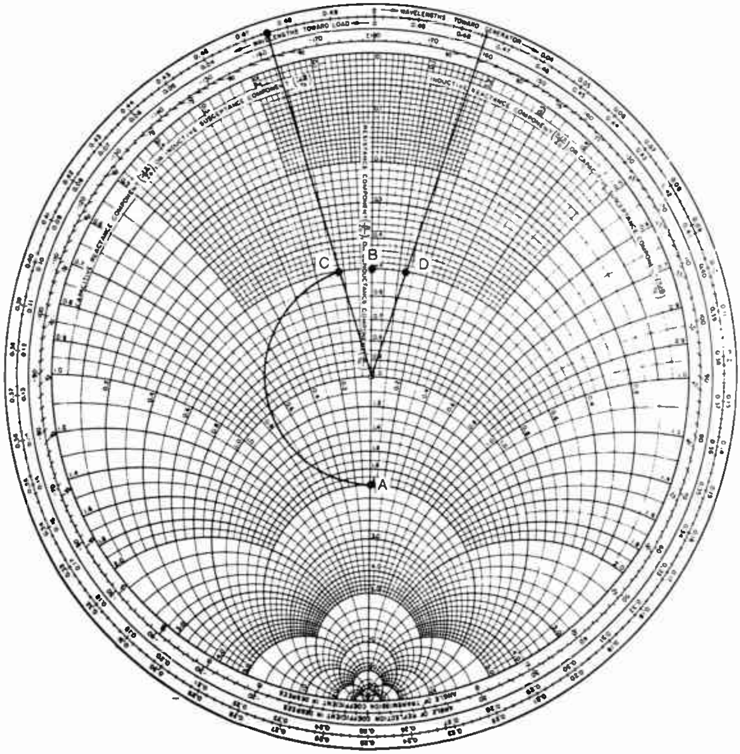


Fig. 4-23. Analysis of line shown in Fig. 4-22.

above and below the center frequency. At the lower frequency the impedance has a slight capacitive component, and at the higher frequency it has a small inductive component. Interestingly, the standing-wave ratio will not change at all with frequency. All of the impedances lie on the same VSWR circle.

Other uses of the Smith chart include plotting the results of measurements of various instruments so that we can find other parameters. For example, if an instrument gives us the magnitude and angle of the reflection coefficient, we can easily find the impedance or admittance.

Plotting Antenna Impedances

In an earlier example we saw that the standing-wave ratio on a transmission line does not vary with frequency as long as the load impedance remains constant. Unfortunately the impedance of most types of antennas does not remain constant as frequency is changed. Hence one of the factors that must be considered when evaluating the bandwidth of an antenna is the standing-wave ratio on the transmission line and how much it varies as the frequency and load are changed. One easy way to evaluate a situation of this type is to plot antenna impedance as a function of frequency on a Smith chart. The standing-wave ratio that occurs at any frequency can be read by simply drawing a VSWR circle through the impedance at that particular frequency.

Suppose that at the terminals of an antenna we measured the driving-point impedance with an impedance bridge and found the following values at different frequencies:

Resistance	*Reactance*	*Frequency, kHz*	*Normalized Impedance*
980	47	$-j12$	$0.94 - j0.24$
990	47	$-j6$	$0.94 - j0.12$
1000	50	0	$1.0 + j0$
1010	53	$+j4$	$1.06 + j0.08$
1020	57	$+j6$	$1.14 + j0.12$

Here we want to know the VSWR at the various frequencies. Plotting the impedances on a Smith chart, we get the plot

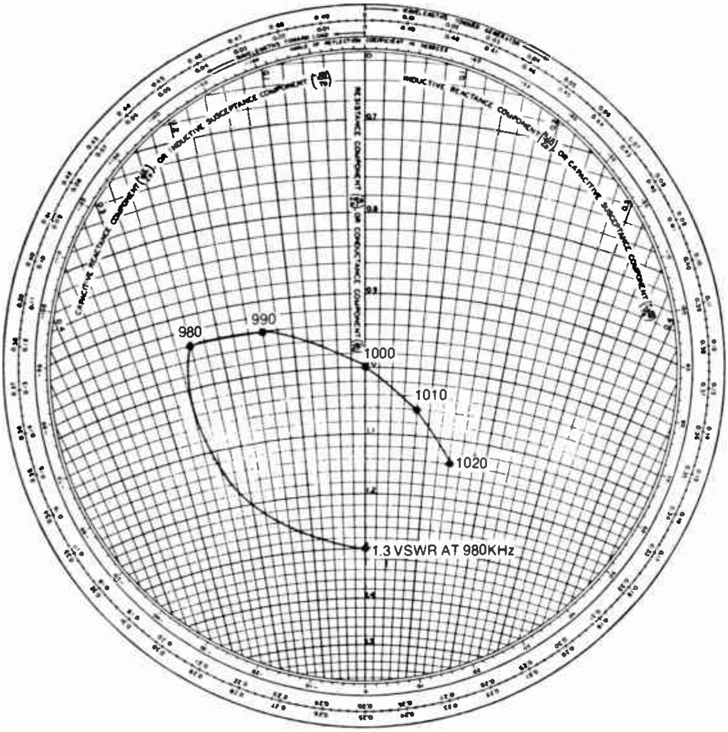


Fig. 4-24. Antenna impedance.

shown in Fig. 4-24. Here we use an expanded chart because it gives us a better view of what is happening. To find the standing-wave ratio at any frequency, we merely draw a VSWR circle through the impedance at that particular frequency. In Fig. 4-24 we see that the greatest standing-wave ratio is obtained at 980 kHz, where it is about 1.3.

In conclusion, the Smith chart is of great practical value because it not only simplifies the computation of transmission-line parameters, it also gives a great deal of insight into the behavior of transmission lines and related components. For a broadcast engineer not already familiar with the chart, it is well worthwhile to get a pad of charts and start plotting impedances and admittances.

Chapter 5

Standard Broadcast Antennas

The history of AM broadcasting in the United States closely parallels the history of broadcast antennas. In the early days of broadcasting, the coverage of a station was determined almost entirely by the transmitter power. At first it looked as if the broadcast band was to be nothing but the scene of a power race, with a few extremely high-powered stations dominating the band. Broadcasting as we know it today is the result of developments in antennas.

The AM broadcast band—or *standard* broadcast band, as it is called in the FCC Rules—consists of 107 channels, between 535 and 1605 kHz. Carrier frequencies in this band are spaced at 10 kHz intervals. Stations are allocated these carrier frequencies in accordance with a system that allows service to almost every section of the country with a minimum of interference between stations. The allocation methods depend heavily on the use of directional antennas.

The history of the directional antenna as applied to AM broadcasting is interesting. To fully appreciate it, one must have some idea of the pandemonium that once prevailed on the broadcast band.

It isn't easy to tell just when the first commercial broadcast station went on the air. The first license to broadcast on a regular basis was issued in 1921 to WBZ, then located in Springfield, Massachusetts. However, many broadcast stations were operating on an experimental basis long before then. Certainly station KDKA in Pittsburgh, Pennsylvania, was one of the earliest.

By 1923 several hundred stations were on the air. At that time only two frequencies, 750 and 833 kHz, were allocated for broadcasting, and the interference was as severe as any of the "pileups" found on crowded ham bands today. Late in 1924, 96 channels, between 550 and 1500 kHz, were set aside for broadcasting, and stations were assigned to particular channels. The earliest attempt to minimize inteference was to assign the same carrier frequencies to stations that were geographically as far apart as possible and, when necessary, to require them to share broadcast time. Old-timers will remember when radio-program listings in the newspapers carried the notation "silent night," indicating that a station would not broadcast on a particular night. The situation was far from satisfactory, but was the best that it could be with the state of the art at that time.

The breakthrough that made broadcasting as we know it today possible was the development in 1921 of the AM directional antenna by Dr. Raymond M. Willmotte and Commander T. A. M. Craven, later an FCC Commissioner. The situation that led to the development involved a case of interference to signals from WTMJ in Milwaukee from WFLA-WSUN in Clearwater, Florida. Both stations operated on a carrier frequency of 620 kHz. A battle ensued, with MTMJ trying to force the Florida station off the air. Some idea of the state of knowledge of propagation at the time can be gathered from the fact that one theory advanced to explain the interference was that the signal from the Florida station traveled across the Gulf of Mexico and then up the Mississippi river valley to Milwaukee—even though this isn't a direct path.

At the time, few broadcast engineers felt that the null from a directional-antenna system could be used to protect the service area of another station from interference. Dr. Willmotte believed that it could, and proceeded to prove it, thus starting the use of directional antennas in broadcasting.

The first AM directional array consisted of two guyed vertical towers, spaced 90° apart and fed through networks that were adjusted without the aid of modern impedance bridges. The results surprised not only the government inspectors but the developers themselves. Dr. Wilmotte reported that when the antenna was adjusted for minimum radiation in the direction of Milwaukee, the radio inspector at Atlanta, Georgia, which is in a direct line between the Florida and Wisconsin cities, wanted to know why the station was off the air without permission.

The directional-antenna system, which is now a familiar sight across the country, became an essential part of the answer to problems of interference between broadcast stations. The allocation of frequencies to broadcast stations depends heavily on its use.

SERVICE AREAS AND CLASSES OF CHANNELS

The FCC recognizes three different types of service area of standard broadcast stations. The *primary service area* is the area where the groundwave from the transmitting antenna is not subject to objectionable fading. The *secondary service area* is the area served by the skywave from the antenna, where there is no objectionable interference. Skywave signals in the secondary service area are, however, subject to fading. The *intermittent service area* lies between the primary and secondary service area. It is the area where the groundwave is received but is subject to fading. Interference may also be present in the intermittent service area.

Three classes of broadcast channels have been established in North America. A *clear channel*, in spite of its name, is not a channel that has only one station assigned to it. It is a channel on which one or more high-powered stations serve wide areas. All of the primary service areas of these stations, and all or a substantial portion of their secondary service areas, are cleared of objectionable interference.

A *regional channel* is one on which several stations with powers of not more than 5 kW operate. Interference between stations is controlled by limiting the contours of their primary service areas.

A *local channel* is one on which several low-powered stations provide service to local communities. Stations operating on local channels operate with a power of not more than 1 kW in the daytime and 250W at nighttime. The primary service areas of these stations are limited by interference considerations.

TIMES OF OPERATION

There are several definitions of times of operation of broadcast stations that must be clarified. *Daytime* and *nighttime* refer to the time between local sunrise and local sunset, and vice versa, but local sunrise is *not* the time that the sun appears over the horizon. Rather the official time of local sunrise, as far as operating rules are concerned, is specified in the station license for each month of the year (and similarly for local sunset).

The *broadcast day* is *not* the same as daytime. It is the period of time between local sunrise, as specified in the station license, and midnight local time.

The *experimental period* is the time period between midnight local time and local sunrise as specified in the station license. During this period any standard broadcast station can broadcast experimentally for purposes of testing and maintenance. The authorized frequency and power may be used for these purposes. However, these experimental broadcasts must not cause objectionable interference to stations that maintain a regular program schedule during this period. No station that is unauthorized to do so may broadcast programs during the experimental period.

STANDARDS OF ALLOCATION

Coverage of the entire country with standard broadcast signals is possible only because of a rather complex system of frequency allocation. Many different broadcast stations operate on the same frequency, and interference between stations is held to a minimum by specifying the maximum and minimum field strength that a particular station can radiate to any particular area. This control is accomplished by limiting the transmitter power and shaping the antenna pattern of each station. The transmitter powers allocated to standard broadcast stations are: 250W, 500W, 1 kW, 2.5 kW, 10 kW, 25 kW, and 50 kW.

For purposes of allocation, stations are divided into four classes and several subclasses. The class of station, authorized operating power, and type of antenna pattern are based not only on the broadcast requirements of the United States but also on international agreements with Canada and Mexico. The object is to provide good broadcast coverage to all areas, while minimizing interference between stations.

Class I stations are dominant stations that operate on clear channels, usually with an operating power of 50 kW, and never less than 10 kW. These stations provide primary and secondary service over a wide area and at long distances from the station. Their primary service areas are cleared of objectionable interference, both on their operating frequency and on adjacent channels. Their secondary service areas are cleared of objectionable interference on their operating frequencies, but not on adjacent channels.

The United States has class I priority on 45 clear channels. Canada and Mexico have their own class I priorities, some of which are shared with the United States. Only one or two class I stations operate on each clear channel.

A *class II* station is a secondary station operating on a clear channel with an operating power of between 250W and 50 kW. Class II stations serve population centers and the adjacent rural areas. They are operated so as not to cause interference with the service areas of dominant stations operating on the same clear channel. There are 29 clear channels on which class II stations may operate.

Class III stations share regional channels with several similar stations, each serving a population center and the surrounding rural areas. Class III stations operate with a power of between 500W and 5 kW. There are 41 regional channels, with more than 2000 class III stations.

A *class IV* station operates on a local channel to provide service to a local area. The operating power is not more than 1 kW during the day and 250W at night. There are six local channels, with 150 or more class IV stations on each channel.

PATTERNS, CONTOURS, AND FIELD INTENSITY

To provide primary service to an area, a station must provide a signal that is strong enough to overcome the manmade noise that might be encountered in the area. The FCC Rules specify the following minimum field intensities for various types of service areas:

Area	*Groundwave Field Intensity*
City business or factory areas	10 to 50 mV/m
City residential areas	2 to 10 mV/m
Rural—all areas during winter, northern areas during summer	0.1 to 0.5 mV/m
Southern areas during summer	0.25 to 1.0 mV/m

These values are based on the absence of fading and interference from other broadcast stations. No real standards of atmospheric or manmade noise have been established, because no uniform measurements are available. The FCC has, however, published a list of signal strengths that are considered satisfactory for overcoming manmade noise in towns of various sizes. The field intensities are:

Population	*Groundwave Field Intensity*
Up to 2500	0.5 mV/m
2500 to 10,000	2.0 mV/m
10,000 and up	Values given in the preceding paragraph

In addition to the requirement for providing adequate field intensity in its service areas, a broadcast station must not radiate interfering signals into the service areas of other stations. This spatial distribution of the signals from broadcast stations is controlled by the design of the antenna system. To specify the field intensity in various directions from an antenna, we need some method of describing how the field intensity varies from one direction to another. This information may be tabulated or given in the form of a graph. Two types of graphs are commonly used to specify the directional characteristics of broadcast antennas—*antenna patterns* and *field-intensity contours*. The two look very much alike and should not be confused.

Antenna Pattern

First let's look at the antenna pattern. To simplify things, we will assume that the antenna is located over a flat, perfectly conducting earth. Suppose that an engineer takes a field-intensity meter and walks around the antenna at some fixed radial distance, say, one mile and as he walks, he periodically reads the field-intensity meter and records its indication. After walking completely around the antenna, he might have a series of measurements as shown in Fig. 5-1.

This information can be used to plot an antenna pattern on a circular chart, with the radial distance from the center

BEARING FROM NORTH, DEGREES	FIELD INTENSITY, mV/m	BEARING FROM NORTH, DEGREES	FIELD INTENSITY, mV/m	BEARING FROM NORTH, DEGREES	FIELD INTENSITY, mV/m
0	500	120	480	240	502
15	504	135	460	255	500
30	510	150	275	270	501
45	498	165	390	285	487
60	500	180	485	300	475
75	502	195	498	315	430
90	509	210	502	330	465
105	501	225	510	345	502

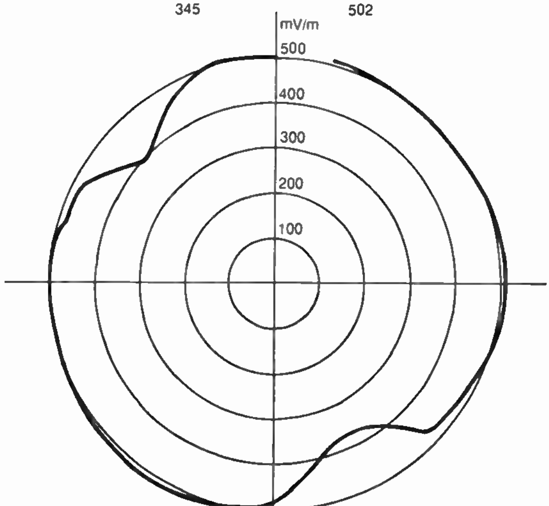


Fig. 5-1. Antenna pattern for data shown.

representing field intensity and the angular scale indicating the bearing at which each measurement was made, as in Fig. 5-1. This particular plot would be the pattern of our antenna at a distance of one mile. If the power of the transmitter was increased, the entire pattern would expand, but the shape would be the same. If the transmitter power was reduced, the entire pattern would shrink, but again the shape would not change. It is customary, when plotting a pattern, to specify the power that is being transmitted.

Although this example illustrates the meaning of an antenna pattern, the assumption was made that the earth in the vicinity of the antenna was a perfect conductor. In the real world this is not true, and we could not determine the actual field intensity at one mile from an antenna by this method. We shall consider the proper method of doing this later.

The portions of the pattern where the signal is strong are called *lobes* of the pattern. If there is one predominating lobe, it is usually called the *major* lobe. Smaller lobes are called *minor* lobes. Those portions of the pattern where the signal is reduced are called *nulls*, or *minima*. Strictly speaking, the term *null* refers to portions where the signal strength approaches zero and *minima* refers to those places where it is merely reduced. However, the use of the term *null* to refer to a bearing where the signal strength is reduced is so widespread that we follow the practice in this book.

Field-Intensity Contour

Another way to describe the directional characteristics of an antenna is to plot a field-intensity contour. Suppose once again that a broadcast engineer, equipped with a field-intensity meter, sets out to measure the directional characteristics of his antenna. Once again he walks completely around the antenna, but this time he does not maintain a constant radial distance from the antenna. Instead, he walks toward or away from the antenna until he obtains a certain indication on his field-intensity meter, say, 1 V/m. Suppose that each time he makes a measurement, he records the angular bearing from true north and the radial distance from the antenna at which the indication of the field-intensity meter is 1 V/m. His measurement record would then appear as in Fig. 5-2.

Now, if he plotted this information on a circular graph, he would obtain the graph shown in Fig. 5-2. This plot is called the *one-volt-per-meter contour* of the antenna. Here again, if the transmitter power was increased, the contour would expand, but its shape would not change. Likewise, if the transmitter power was reduced, the contour would shrink without changing shape. To be meaningful, the contour should also specify the transmitter power.

The contour is important in that the service area of a station is protected to a minimum signal contour against interference from other stations. Another use of the contour is in assigning responsibility for cases of interference to listeners' receivers. The FCC Rules provide that any standard

BEARING FROM NORTH. (DEGREES)	DISTANCE FROM ANTENNA MILES	BEARING FROM NORTH. (DEGREES)	DISTANCE FROM ANTENNA. MILES
0	0 32	120	0 30
20	0 33	140	0 28
40	0 34	160	0 22
60	0 31	180	0 20
80	0 28	200	0 22
100	0 29	220	0 20

BEARING FROM NORTH. (DEGREES)	DISTANCE FROM ANTENNA MILES
240	0 20
260	0 20
280	0 25
300	0 26
320	0 26
340	0 29

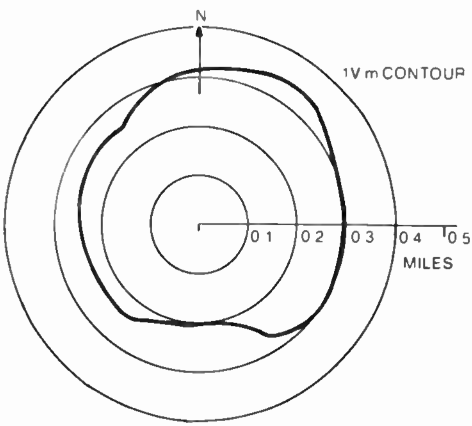


Fig. 5-2. Field-intensity contour.

broadcast station is responsible for adjusting all cases of interference to listeners within its one-volt-per-meter contour.

We have now seen examples of both patterns and contours. They are two different ways of describing the directional characteristics of antennas. The pattern is a plot of field intensity at a fixed distance from the antenna at various angles. The contour is a plot of the distance from the antenna to a point of given field intensity at various angles. The distinction should be kept clearly in mind.

Vertical Pattern

The signal serving the primary service area of a station is propagated along the surface of the earth and is called the *groundwave*. It is specified in terms of a pattern calculated along the surface of the earth. Signals are also propagated by the *skywave*, which is reflected from the ionosphere back to the earth. Skywave propagation provides coverage to secondary service areas of some stations and is a potential source of interference to other co-channel and adjacent-channel stations.

The amount of signal that an antenna radiates at various angles above the horizon is specified by the *vertical-radiation pattern*. The vertical pattern, like the horizontal pattern, is usually plotted on polar graph paper. The radial scale is calibrated in field intensity, usually at a distance of one mile from the antenna. The angular scale is simply the angle from the horizon. Figure 5-3 shows the vertical-radiation pattern of

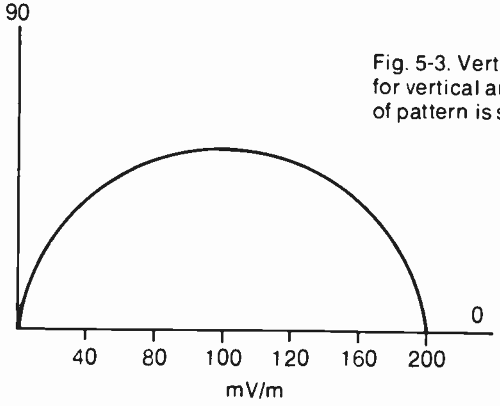


Fig. 5-3. Vertical-radiation pattern for vertical antenna (only one side of pattern is shown).

a vertical antenna that is 1/4 wavelength in height. Note that the radiation is greatest along the surface of the earth but that there is still appreciable radiation at higher elevation angles. Usually, radiation above 60° is small, and not subject to propagation over long paths. It is the radiation at elevation angles below 60° that is responsible for both coverage to secondary service areas and interference to other stations.

TYPICAL STANDARD BROADCAST ANTENNAS

There are two types of antenna systems used in standard broadcast stations: nondirectional antennas, which consist of a single vertical tower, and directional antenna arrays, which have two or more towers.

A nondirectional antenna consists essentially of a vertical tower, which radiates the signal; a network, which matches the impedance of the antenna to the characteristic impedance of the transmission line; and the transmission line itself. A typical system is shown in Fig. 5-4. Most broadcast transmitters are designed to work into the characteristic impedance of a transmission line, so no matching network is required between the transmitter and the line.

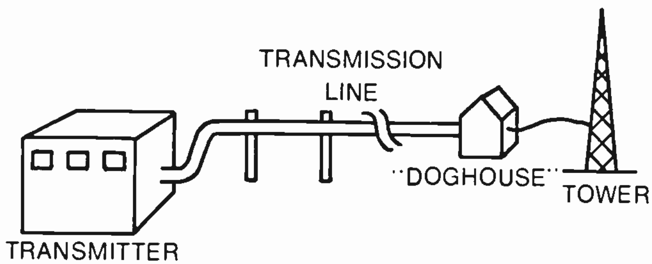


Fig. 5-4. A typical nondirectional antenna system for AM broadcasting.

Accessories

The driving-point impedance seen at the base of a tower is never the same as the characteristic impedance of a transmission line. Thus a matching network is required. This network is located close to the base of the tower in a shelter that is commonly known as a "doghouse." Sometimes this unit is called a *line-tuning unit*, or LTU. In a nondirectional antenna the phase of the current feeding the antenna is not important, so the phase characteristics of the matching network are not critical. About the only critical requirements of the matching network are low losses and a bandwidth adequate to transmit all of the sideband power.

The transmission line is almost always a coaxial cable. A few open-wire lines are extant, but these are rapidly being replaced by coaxial cables. The requirements for the cable are low losses and adequate power rating to handle the licensed transmitter power.

Besides the essential elements of the antenna and feeder, certain accessories are required. One of these is a tower-lighting system. The FCC and FAA rules require that all broadcast towers be lighted at night. If a tower is shunt fed, as in Fig. 5-5A, there is no problem; the lighting wires are simply run up along the side of the tower. Unfortunately, shunt feeding of broadcast towers poses some problems, and it is not commonly used. The usual broadcast-antenna tower has a base insulator (Fig. 5-5B), and the base of the tower is not at ground potential. If the wires carrying power to the tower lights in B were run as in A, they would effectively short-circuit the signal at the base of the tower. It is necessary, therefore, to run the lighting power through some sort of

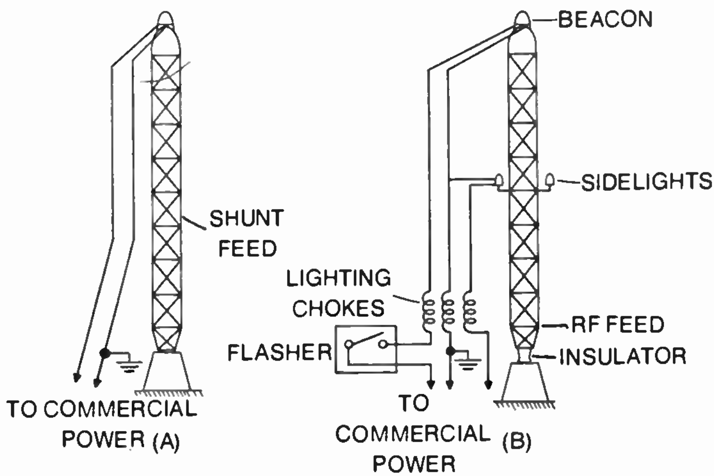


Fig. 5-5. Tower lighting arrangements for shunt- and series-fed AM antennas.

arrangement that will present a low impedance to the 60 Hz power but a high impedance to the RF signal.

A typical solution is to include isolation chokes, as shown in Fig. 5-5B. Another solution is the Austin transformer (Fig. 5-6). The two coils of this transformer are magnetically coupled so that the 60 Hz power will pass between them. They are physically separated by a great enough distance that the capacitive coupling between the primary and secondary will be small, and little signal power will be coupled through the transformer.

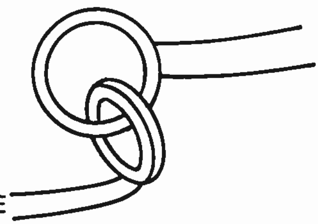


Fig. 5-6. Austin transformer for tower lighting.

Another essential part of the antenna system is a lightning-protection device. This usually takes the form of an air gap that will break down when the tower becomes charged.

Parts of the Directional Antenna

A directional-antenna system contains everything that is in the nondirectional antenna, and more. Figure 5-7 shows a block diagram of a 3-tower directional-antenna system. As in the nondirectional antenna, it is necessary to match the driving-point impedance at the base of each tower to the characteristic impedance of the transmission line. But in the directional antenna the relative phase of the current fed to each tower is critical, so an additional requirement is imposed on the matching network. It must not only perform the required impedance transformation, but it must do so with a predetermined amount of phase shift.

At the sending end of the transmission lines feeding each of the towers, the signal must be divided so that each tower carries the proper percentage of the total current at the proper phase. Controls are provided to adjust the magnitude and phase of each of the tower currents. The equipment used for

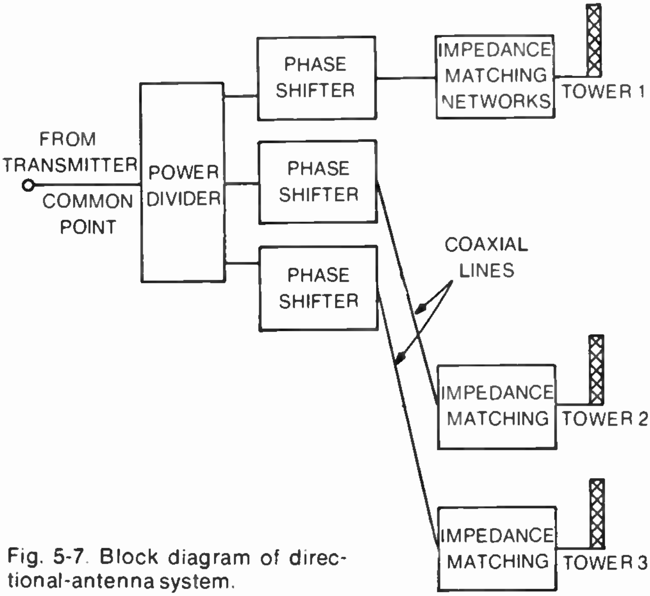


Fig. 5-7. Block diagram of directional-antenna system.

this purpose consists of phase-shifting networks and power dividers. This equipment is usually referred to collectively as the *phasor*.

The point in the system just before the power from the transmitter branches out is called the *common point* of the system. Here the impedance and current are measured to determine the power radiated by the system.

One subsystem of a directional-antenna system that is not used in a nondirectional antenna is the *antenna-monitoring* system, formerly called the *phase-monitoring* system. This system is used to measure the amplitude and phase of the currents in each of the antenna towers to ensure that the pattern stays within its licensed limits. Small samples of the tower currents are picked up by *sampling loops* mounted on the side of each tower. The signals from the loops are carried back to the transmitter building, through coaxial cables to the antenna monitor (formerly called the *phase* monitor). The antenna monitor indicates the ratios of the currents in the various towers of the array, as well as their phase angles.

In general, the power delivered by the transmitter to the antenna system is given by the equation

$$P = I^2 R$$

where P = power from transmitter
R = resistance measured at common point of system
I = unmodulated RF current measured at common point

To make an allowance for the additional losses that are unavoidable in a directional-antenna system, the FCC Rules state that the power shall be determined from the relationship

$$P = I^2 Ra$$

where the constant a equals 0.92 in stations where the licensed power is 5 kW or less, or 0.947 where the licensed power is over 5 kW.

Another important part of a broadcast-antenna system is the current-measuring system. The rules require that the current at the base of each tower be measured at regular intervals. This measurement is made by a thermocouple ammeter, called the *base-current* ammeter. To protect the meter from damage that might be caused by lightning surges, it is usually switched out of the circuit except when measurements are actually being made.

PROPAGATION OF STANDARD BROADCAST SIGNALS

At the frequencies used for standard AM broadcasting, there are two primary modes of signal propagation—groundwave and skywave propagation. During daylight, propagation is entirely by means of the groundwave. It is this mode of propagation that provides coverage of the primary service area of the station. Starting about local sunset, signals begin to be propagated by the skywave. In this mode, signals are reflected from the ionosphere back toward the earth. Skywave propagation provides coverage of the secondary service area of a station, if the station has such coverage. It may also cause interference to adjacent channel or cochannel stations.

Figure 5-8 illustrates both modes of propagation. Part A shows that the groundwave is strongest close to the antenna and falls off with distance. If the ground were a perfect conductor, the signal level would decrease linearly with distance from the antenna; but because the ground is not a

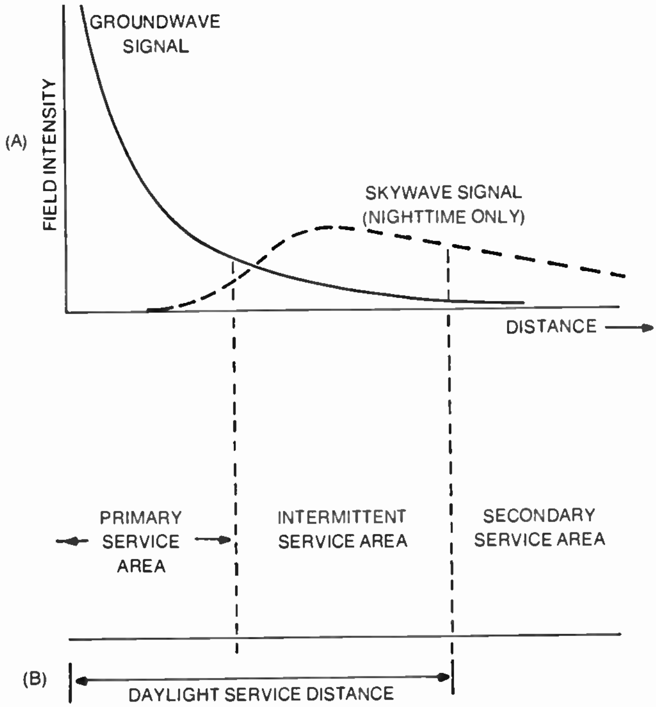


Fig. 5-8. Groundwave and skywave propagation.

perfect conductor, the signal actually falls off more rapidly. During the daylight hours, when there is no skywave, the groundwave signal can be received until it is buried in the noise. At night, when the skywave signal appears, it is greater than the groundwave signal at a distance considerably less than the daytime range of the station.

When the groundwave and skywave signals are nearly the same strength, they alternately reinforce and cancel each other, leading to serious fading of the signal at the receiver. The area where this occurs is called the *intermittent service* area (Fig. 5-8B). The point where fading starts to become objectionable—at the so-called *fading wall*—marks the outer edge of the primary service area. One of the major considerations in broadcast-antenna design is to push back the fading wall. This is done by confining as much of the radiation as practicable to low elevation angles so as to reduce skywave radiation.

Chapter 6
Vertical Antennas

The vertical antenna is the exclusive radiating element in standard broadcasting. It is well suited to broadcast use. The vertical antenna's radiation is uniform along the surface of the earth, and because of its low angle of radiation, most of the energy is concentrated in the groundwave, which provides primary service. The slightness of radiation at high vertical angles minimizes skywave interference between stations at night.

The vertically polarized signal from a vertical antenna suffers much less loss from low ground conductivity than a horizontally polarized signal does. This increases the area of groundwave coverage. Vertical antennas are also well suited for use as elements of directional antennas for standard broadcasting. Since a vertical antenna operating alone has a circular pattern in the horizontal plane, the design of arrays that must produce complex patterns is simplified.

In standard broadcast work we are interested in four properties of the vertical antenna:

1. The amount of radiation, or signal strength, along the surface of the earth
2. The distribution of energy at vertical angles above the earth

3. The driving-point impedance, that is, the impedance seen across the terminals where energy is fed to the antenna (usually across the base insulator)
4. The losses associated with the antenna

In studying the vertical antenna, we will first make some assumptions that, although not always true, considerably simplify our analysis. We will then modify our results so that they can be applied to practical antennas.

BASIC PRINCIPLES OF VERTICAL ANTENNAS

In our study of radiation in an earlier chapter, we worked with antenna elements that were dipoles. Much of our work was with the half-wave dipole. We didn't consider the effect of the ground as far as the antenna itself was concerned. The ground however, is an essential part of the vertical antenna and must be considered in all practical work. For now, let us think of the ground as a flat, perfectly conducting plane. We will later consider the effect of the finite conductivity of the ground.

Our perfectly conducting ground can be thought of as being a large mirror, as far as radio energy is concerned. Thus the ground reflects any energy that is radiated downward from an antenna mounted above it. If a vertical quarter-wave antenna is mounted on the surface of the earth, the reflection will make it "look like" a half-wave dipole, as shown in Fig. 6-1. The ground takes the place of the "missing" 1/4

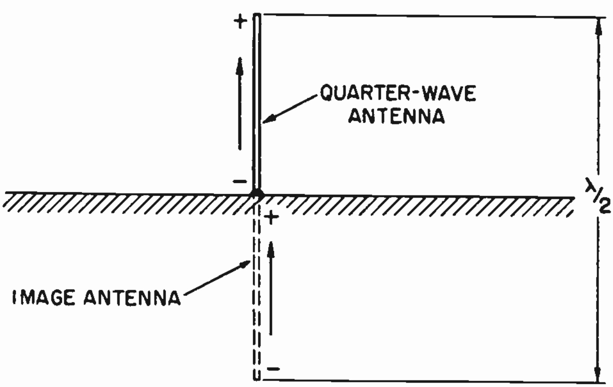


Fig. 6-1. Image antenna.

wavelength of antenna, and reflection from the ground supplies the energy that would be radiated by the "missing" section.

Understand that the image antenna doesn't actually exist. If there were a tunnel or cave under the antenna where measurements could be made, we wouldn't find the voltage or current shown in Fig. 6-1. The way in which an image antenna is formed by reflection can be understood by considering what happens when a flashlight is directed into a mirror. In Fig 6-2 the flashlight is pointed in such a way that the light shines directly into the observer's eyes. The flashlight is actually pointed away from the observer, but the image flashlight in the mirror is pointed directly at him. The effect is the same as if an actual flashlight was located behind the mirror site with the mirror removed. This is directly analogous to the creation of an image antenna be reflection from the ground. Just as the direction in which the flashlight is pointing is reversed by the reflection, so the polarity of the charge on the image antenna will be reversed.

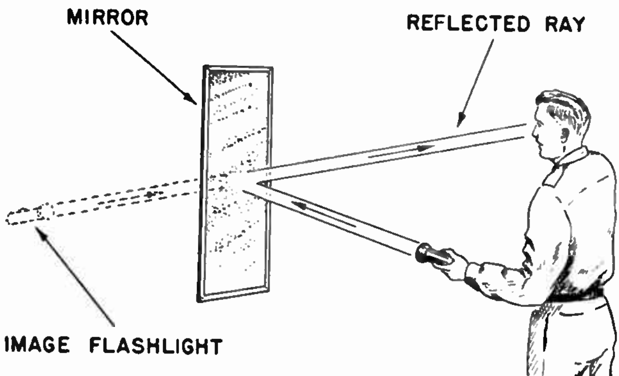


Fig. 6-2. Explanation of image-antenna principle.

Current Distribution

The pattern of the vertical antenna, and its driving-point impedance are affected to some extent by the way the current is distributed along the antenna. A rigorous analysis of current distribution is quite complicated. Fortunately, such an analysis isn't necessary. We can make a few assumptions that will be adequate for most purposes.

Figure 6-3 shows a quarter-wave vertical antenna. If the antenna were infinitely thin, the current distribution would be sinusoidal. The current would be zero at the top, where there is no place for it to flow, and maximum at the base. In the same way, the voltage would be maximum at the top, where no current flows, and minimum at the base. In actual antennas the current distribution isn't exactly sinusoidal, but it is nearly so. Later we will see how the departure of the current from a sinusoidal distribution affects the radiation pattern and the driving-point impedance.

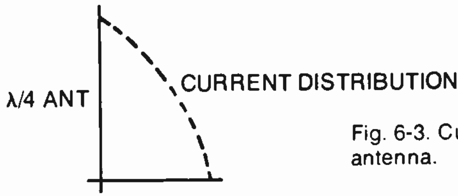


Fig. 6-3. Current on quarter-wave antenna.

Figure 6-4 shows a vertical antenna that is arbitrarily longer than 1/4 wavelength. To find the current distribution, we can start at the top, where the current is zero and the voltage maximum. The current will increase and the voltage will decrease as we proceed down the tower.

If the antenna were ideal (without losses), the current would decrease as shown by the solid line in Fig. 6-4A. Actually, in a practical antenna, it will be more as shown by the dashed line. Even then, there will be a great deal of variation from one tower to another, depending on the physical configuration of the tower. Similarly, if there were no losses, the phase distribution would be as shown by the solid line in Fig. 6-4B. The current would have one phase above the quarter-wave point and another phase below it. That is, the current would be flowing in one direction in the top part of the antenna and in the other direction in the bottom part. In an actual tower the current direction does reverse at about the quarter-wave point, but the transition is more gradual, as shown by the dashed line in Fig. 6-4B. We will have many occasions to consider the current distribution on a tower, because many of the properties of an antenna are intimately related to its current distribution.

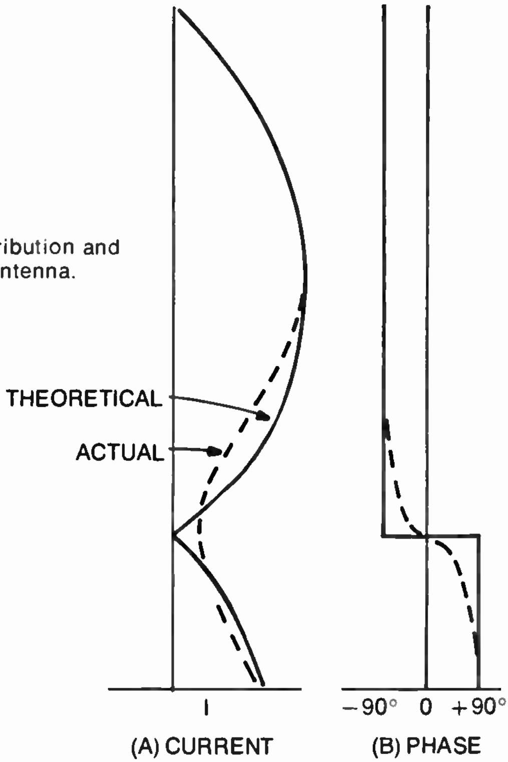


Fig. 6-4. Current distribution and phase on a practical antenna.

Radiation at Vertical Angles

A 3-dimensional view of the radiation pattern of a vertical tower is shown in Fig. 6-5A. As long as the ground conductivity is uniform around the antenna, it will radiate equally well in all directions in a horizontal plane; that is, the pattern along the ground will be a circle (Fig. 6-5B). This pattern is the same as would be seen in looking down on the doughnut-shaped pattern of Fig. 6-5A from the top.

Although in broadcast work we are primarily interested in how well an antenna radiates energy along the surface of the earth, we cannot ignore its vertical pattern or how well it radiates energy at vertical angles above the horizon. Radiation at vertical angles is of interest for three reasons:

1. The energy that is radiated at vertical angles is not available for coverage of the primary service area of the station.

2. Radiation at vertical angles provides coverage of the secondary service area of the station.
3. Radiation at vertical angles may cause interference to cochannel or adjacent-channel stations at night.

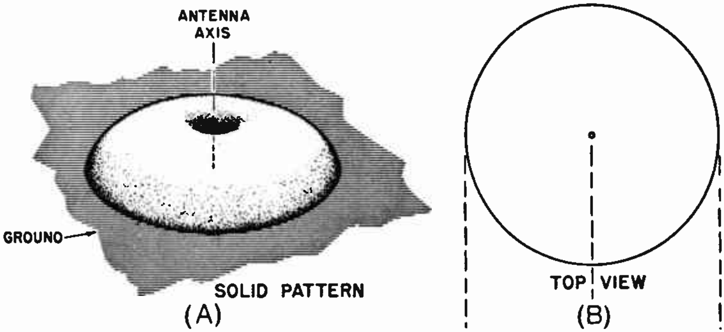


Fig. 6-5. Radiation pattern of a vertical antenna.

EFFECTIVE FIELD INTENSITY AT ONE MILE

All AM station allocations and antenna designs are based on the *effective* (or *inverse* or *unattenuated*) field intensity at one mile from the antenna E_{mi}. (This is sometimes called simply the *effective field.*) This is the field intensity that would be produced if the antenna were located over a perfectly conducting earth. It definitely is not the field intensity that we would actually measure at a distance of one mile from the antenna.

In making preliminary computations, we can use this theoretical effective field intensity. After a station is installed, the actual effective field intensity is determined from actual field-intensity measurements. Many different measurements are made to determine this value, so it is quite representative of the actual behavior of an antenna.

Ideal Hemispherical Radiator

The vertical-radiation pattern of an antenna depends on its height. As a reference we often use the theoretical pattern of an ideal hemispherical radiator that is mounted on the surface of the earth. Of course, such an antenna is not realizable physically, but if it were, it would have a semicircular pattern in the vertical plane (Fig. 6-6). It is easy to calculate what the field intensity of such an antenna would be at a given distance from the antenna; thus it makes a convenient reference.

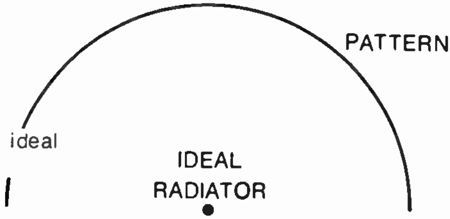


Fig. 6-6. Pattern of ideal hemispherical radiator.

Calculating Effective Field of Ideal Radiator

The area of a sphere is given by

$$A = 4\pi r^2$$

where A is the area, and r is the radius. The area of a hemisphere is half this value; that is,

$$A = 2\pi r^2$$

Now, we want to find the field intensity from our ideal hemispherical radiator at a distance of one mile. Since one mile is equal to 1609m, the area in square meters of a hemisphere with a radius of one mile is

$$A = 2\pi \times (1609)^2 = 16{,}266{,}419\ m^2$$

Let's assume that our ideal radiator is radiating 1 kW of power with 100% efficiency. The power density p at the surface of the sphere with a one-mile radius will be

$$p = \frac{1000}{16{,}266{,}419^2}\ 0.00006\ W/m^2$$

The relationship between radiated power and field intensity at some distance from the antenna is

$$E = \sqrt{p \times 377}$$

where p = power density in watts per square meter
E = field intensity in volts per meter
377 ohms = impedance of free space

Therefore the field intensity from our ideal radiator *at one mile* is

$$E = \sqrt{0.00006 \times 377} = 0.152\ V/m$$

This means that the field intensity along the earth at a distance of one mile from an ideal, *infinitely short* antenna (uniform hemispheric radiator) would be 152.2 mV/m.

The field intensity at a distance of one mile for antennas of other heights can be found by similar methods. Inasmuch as their patterns are not hemispherical, the mathematical operations for these antennas are more complicated. We will not calculate the field intensities here, but will merely tabulate them.

Figure 6-7 shows four vertical antennas of different heights together with their vertical-radiation patterns. Notice that as the height of the antenna is increased, the vertical-radiation pattern is squashed, so that more energy is radiated along the surface of the earth. The signal intensity increases as the height of the antenna is increased, until a maximum is reached when the antenna height reaches 225°. It is not always practicable to use antennas of this height, because after a height of 180° is reached, a minor lobe starts to form at a vertical angle of about 60° (D in Fig. 6-7). This minor lobe can be desirable in that it may increase coverage of the secondary service area; or it can be undesirable in that it may cause skywave interference to other stations at night.

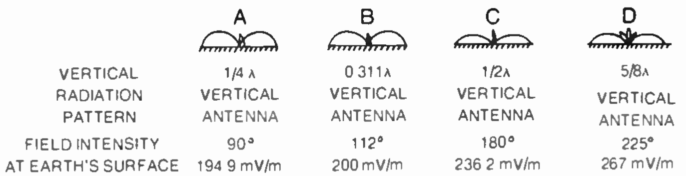


Fig. 6-7. Vertical-radiation patterns for antennas of various heights.

GROUNDWAVE SIGNALS

The primary service area of a standard broadcast station is served by the groundwave. If the earth were a perfect conductor, computing the field intensity of the groundwave signal would be very simple. Figure 6-8 gives the field intensity at one mile from the antenna for various antenna heights, with a radiated power of 1 kW. To find the *unattenuated* field intensity for any other distance or radiated power, we merely have to substitute values into the equation

$$E = \frac{E_0 \sqrt{P}}{d}$$

where d = distance from transmitter in miles
E_0 = intensity of field at one mile, 1 kW
P = radiated power in kilowatts

Suppose, for example, that we have a 5 kW station with a 90° antenna and wish to find the unattenuated field intensity at 2 miles from the antenna. From Fig. 6-8 we find the value of E_0 to be about 195 mV/m. The other parameters are $d = 2$ and $P = 5$. Substituting these values into the above equation gives us

$$E = \frac{195\sqrt{5}}{2} = 218 \text{ mV/m}$$

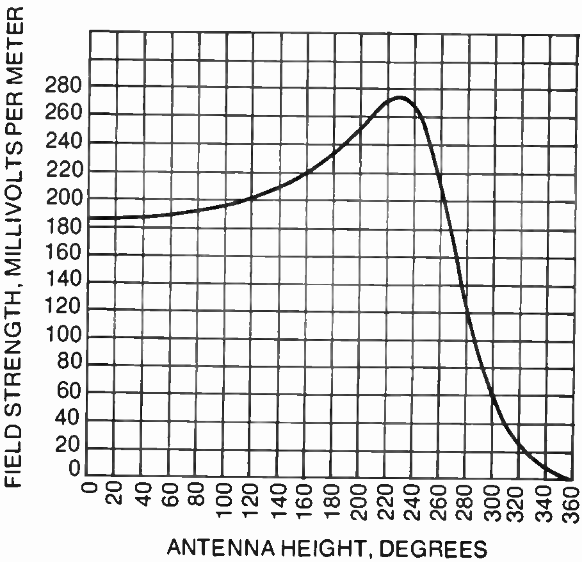


Fig. 6-8. Field intensity at one mile and 1 kW versus antenna height.

Figure 6-9 shows a plot of the vertical-radiation characteristics of vertical antennas taken directly from the FCC Rules. The curves at the left represent field intensities in millivolts per meter for a radiated power of 1 kW. They give the field intensity at one mile along the surface of the earth and at all vertical angles. Inasmuch as there is no radiation at all from the top of a vertical antenna, the curves all go to zero at 90°.

The curves at the right of Fig. 6-9 are apt to be confusing. The radiated power is not specified and is not assumed to be constant. The radiated power for each antenna height is adjusted to produce a field intensity along the surface of the earth of 100 mV/m at one mile from the antenna. Thus the curve for each antenna height shows what the field intensity will be at a radius of one mile for all vertical angles if the field intensity along the surface of the earth is 100 mV/m. Both sets of curves of Fig. 6-9 are useful for calculating the field intensity that the antenna will produce at great distances from the antenna by skywave propagation.

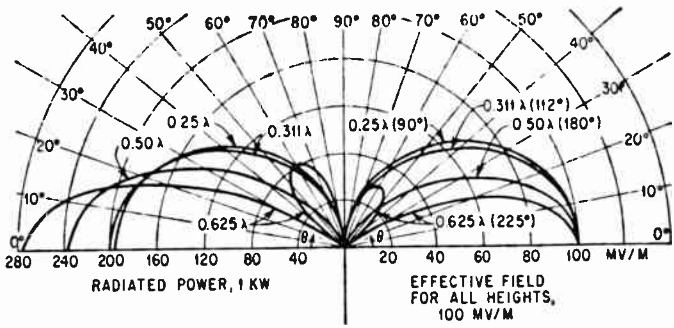


Fig. 6-9. FCC vertical-radiation characteristics.

Earlier we computed a field intensity for a distance of 2 miles from an antenna, assuming the ground was a perfect conductor. The calculation is useful in that it shows the theoretical maximum field intensity, but it doesn't tell us the actual field intensity. The ground is not a perfect conductor; in fact, it is a rather poor conductor that has both resistance and capacitance. The amount of resistance and capacitance depends on the frequency of the signal and the composition of the earth in the region of interest. At standard broadcast frequencies the earth acts as a resistance, and we can ignore the capacitive effects.

It would seem at first glance that the earth is such a large conductor that its resistance shouldn't have much effect on a signal. This isn't true, because the skin effect confines signals to a layer close to the surface of the earth.

Ground Conductivity

The attenuation of radio signals due to the earth's resistance is normally expressed in terms of *conducitivity* rather than resistance. Conductivity is the measure of the ability of a specific material, such as the ground, to conduct electricity. It is therefore different for different materials. It is the reciprocal of *resistivity* ρ, which is given as

$$\rho = \frac{RA}{l}$$

where R is the resistance of a certain specimen of wire or other conductor, A is the area of the specimen, and l is the length. Conductivity σ, then, is given as

$$\sigma = \frac{l}{RA}$$

Conductivity is thus stated as so many mhos per unit length. For the earth it is given as so many mhos or millimhos per meter. If a solid cube of earth, one meter on a side, has a conductance of one mho between opposite faces, it has a conductivity of one mho per meter.

Much of the literature dealing with the conductivity of the earth gives conductivity in *electromagnetic units*. These are units of the cgs electromagnetic system of units, which was formerly widely used in scientific work. To convert from electromagnetic units (emu) of conductivity to millimhos per meter (mhos/m) simply multiply by 10^{14}. For example, the conductivity of sea water is about 5000×10^{14} emu. The conversion is as follows:

$$5000 \times 10^{-14} \text{ emu} \times 10^{14}\ 5000 \text{ mmho per meter}$$

Figure 6-10 shows the ground conductivity for the 48 conterminous states of the U.S. and the lower part of Canada. (The U.S. map is included in the FCC Rules as Fig. R-3, and the Canadian map is available from Canada's Department of Communication.) From such maps we can find the approximate ground conductivity in any given area. Later we will see how to determine the ground conductivity in the primary service area of a station by using data from field-intensity measurements.

Groundwave Field Intensity

Once we know the ground conductivity in an area, we can compute the groundwave field intensity as a function of

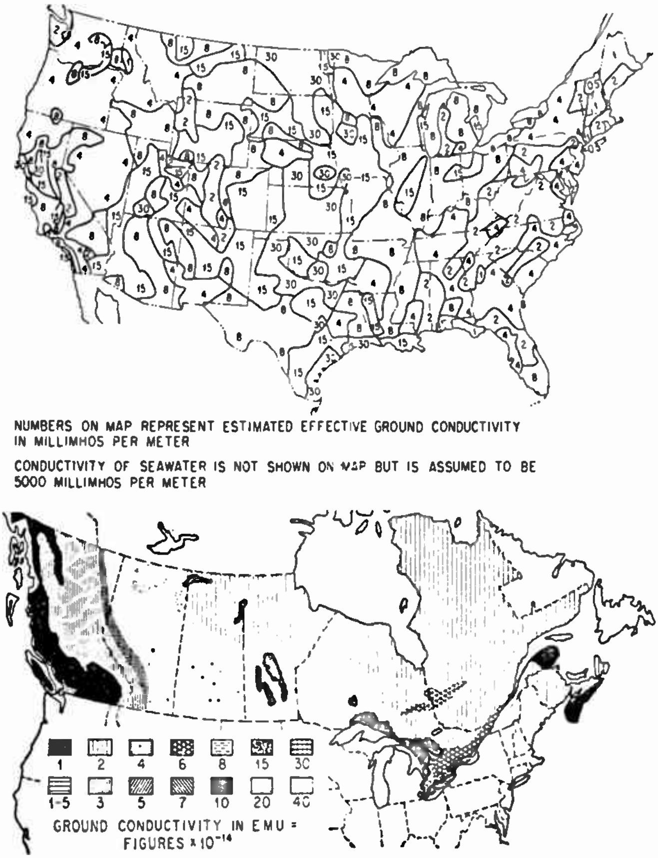


Fig. 6-10. Numbers on map represent estimated effective ground conductivity in millimhos per meter. Conductivity of seawater is not shown on map, but is assumed to be 5000 millimhos per meter.

distance. A common problem in determining whether or not interference will exist between stations is to determine the distance to a given field-intensity contour. The FCC Rules contain 20 graphs giving field intensity as a function of distance and ground conductivity. These graphs cover all of the frequencies in the standard broadcast band.

Figure 6-11 shows the FCC graph covering signals between 970 and 1030 kHz. Note that the curves in this graph are based on an unattenuated field intensity of 100 mV/m at one mile from the antenna. If the antenna does not produce a field intensity of 100 mV/m at one mile, which is likely, it will be necessary to scale the parameters.

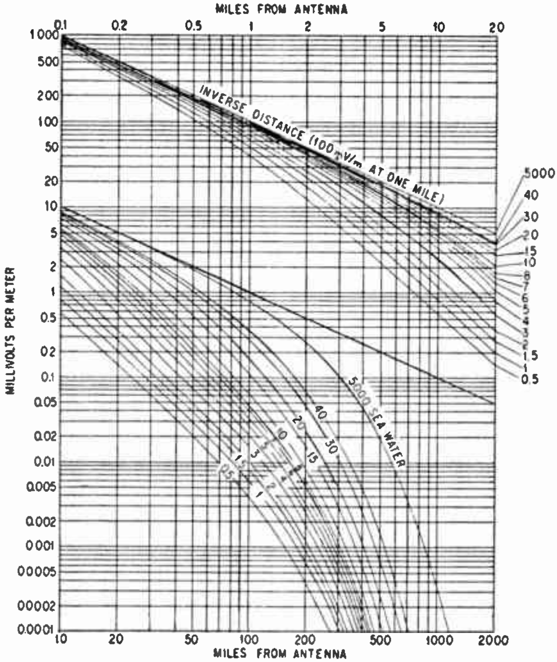


Fig. 6-11. Field-intensity versus distance.

Assume that an antenna produces an unattenuated field intensity at one mile of 100 mV/m and is located in an area where the ground conductivity is 6 mmho/m. Suppose we wish to find the distance to the 500 μV/m (0.5 mV/m) contour. We can read the distance to the 0.5 mV/m contour directly from the chart as 32 miles.

In most cases, the unattenuated field from the antenna at one mile will have some value other than 100 mV/m. Suppose for example, that the field intensity is 175 mV/m. To use the curves, we must multiply the desired contour, 500 μV/m, by the ratio of 100 mV/m to the actual unattenuated field intensity, 175 mV/m.

$$0.5 \frac{100}{175} = 285 \mu\text{V/m}$$

This means that to find the distance to the 500 μV/m contour of our *antenna*, we will have to find the distance to the 285 μV/m contour in the *chart*. We can now read this distance as about 39 miles.

MINIMUM ANTENNA HEIGHT

The FCC station allocations rely heavily on the unattenuated field from an antenna at a distance of one mile from the antenna. Of course, the actual field intensity is less than this value because of the attenuation of the earth; nevertheless, this figure is very useful in calculating actual coverage and interference contours.

The FCC Rules require that any new station, or any station undergoing major modifications, have an antenna system that meets certain minimum standards. Figure 6-12 shows curves from the FCC Rules that specify the minimum acceptable antenna height for each class of standard broadcast station. If these minimum heights are not met, the Rules require the station to submit evidence that minimum field intensities at one mile from the antenna are provided. The requirements are summarized below.

Class IV stations must have an antenna height at least as great as that shown by curve *A* of Fig. 6-12. If the class IV

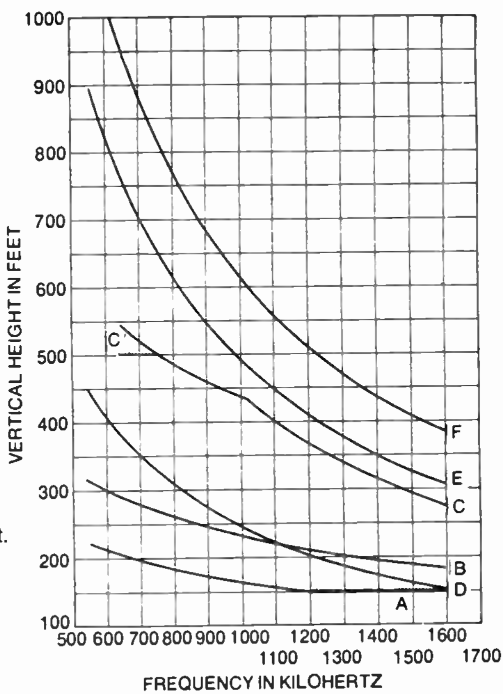


Fig. 6-12. FCC curves for minimum antenna height.

station is assigned to a local channel, it may, in lieu of meeting the antenna-height requirement, submit evidence that the effective field intensity at one mile is at least:

Power	*Field Intensity*
1 kW	150 mV/m
250W	75 mV/m

Class II and class III stations must provide a minimum effective field intensity at one mile from the antenna of 175 mV/m for one kilowatt of radiated power.

Class I stations must provide a minimum effective field intensity at one mile from the antenna of 225 mV/m for one kilowatt of radiated power.

SKYWAVE PROPAGATION

During daylight the groundwave is the only mode of propagation of standard broadcast signals. At night the signals are also propagated by the skywave. In (skywave) propagation the signals that are directed upward above the horizon are refracted (bent) by the *E*-layer of the ionosphere and directed back toward the earth. The action of a signal on reaching the ionosphere is not a simple reflection, but rather a gradual bending until the signal is directed back toward the earth. The process of refraction is quite complex, but by means of a few simplifying assumptions, we can quite easily find the approximate field intensity of skywave signals.

To keep the mathematics simple, we will assume that signals reaching the ionosphere are actually reflected from a virtual height, as shown in Fig. 6-13. Because of the curvature of the earth, the signals leave the earth on an angle θ and hit the virtual reflecting layer at a slightly greater angle Φ. The relationship between the two angles is given by

$$\cos \Phi = \frac{\cos \theta}{1 + h_v / r_e}$$

where θ = elevation angle

Φ = incident angle with reflecting layer

h_v = virtual height of reflecting layer

r_e = vadius of earth (in the same units as h)

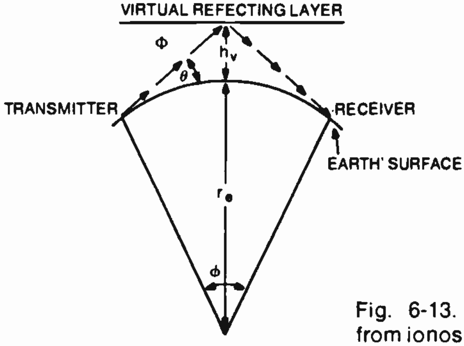


Fig. 6-13. Reflection of signal from ionosphere.

The sum of the interior angles of a triangle is 180°; therefore we can write the equation

$$\phi = 2(\Phi - \theta)$$

where ϕ is the angle at the center of the earth (Fig. 6-13). Now, there is one more principle from geometry that will enable us to find the distance d between the transmitting antenna and the point where the signal returns to the earth: The distance along any great circle on the surface of the earth is given by

$$d = r_e \phi$$

where d is in the same units as r_e, and ϕ is expressed in radians.

The remaining step is to plug numbers into the equations. The radius of the earth is 3960 miles, and the commonly accepted value for the virtual height of the E-layer is 110 kilometers, or 68.35 miles. We now can write the equations

$$d = 138.2\,(\Phi - \theta)$$

$$\cos \Phi = 0.983 \cos \theta$$

where d is in miles, and the two angles are in degrees.

Suppose, for example, that we wish to find the distance between a transmitting antenna and the point on the surface of the earth where a signal will return if it is radiated at an angle of 15°. Using the above equations,

$$\cos \Phi = 0.983 \cos \theta$$
$$= 0.983 \cos 15° = 0.9495$$
$$\Phi = 18.3°$$
$$d = 138.2\,(18.3° - 15°) = 454 \text{ miles}$$

PROBABLE INTENSITY OF SKYWAVE SIGNAL

The amount of ionization at any layer in the ionosphere depends not only on the time of day but also on such things as sunspot cycles. Our field-intensity computations are therefore approximate. For this reason it is customary to express the field intensity of skywave signals in statistical rather than absolute terms. We thus state a field intensity that we might expect to be exceeded part of the time. In station-allocation computations the FCC uses values that might be expected to be exceeded 10% and 50% of the time. Curves are given in the Rules that simplify the computations.

Figure 6-14 shows a graph of the field intensities that might be expected 10% and 50% of the time at various distances from the transmitting antenna when the radiation at the pertinent elevation angle has a field intensity of 100 mV/m at one mile from the antenna. Thus this graph can be used to get a statistical measure of the field intensity of the skywave at any distance from the antenna. The first thing that we must know to use these curves is the pertinent angle of elevation, that is, the angle of radiation or departure. We can get this from Fig. 6-15, which is taken from the FCC Rules.

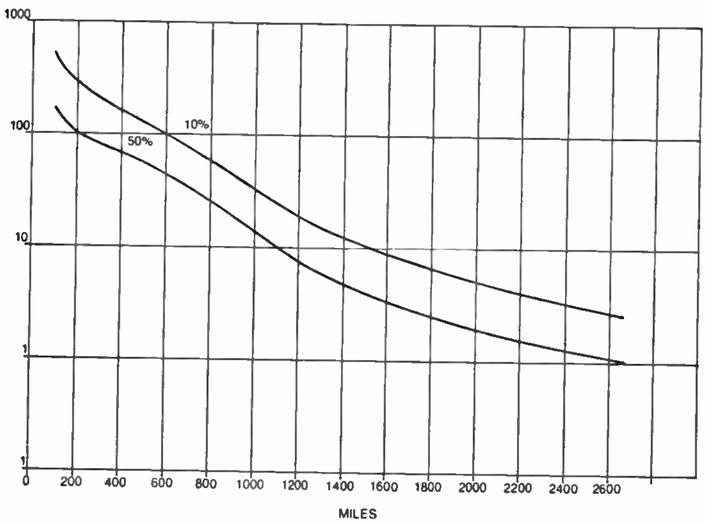


Fig. 6-14. Skywave signals for 10% and 50% of the time. Skywave range for frequencies 540 kHz to 1600 kHz, based on a radiated field of 100 mV/m at one mile at the pertinent vertical angle.

Suppose, that we have a 10 kW transmitter with a 90° antenna and we wish to know the field intensity that we can expect 50% of the time at a distance of 500 miles from the transmitting antenna. From Fig. 6-15 we can see that the signal reaching the earth at this distance must be radiated at an elevation angle of about 11°. Now we need to know the field intensity that our antenna radiates at this angle. We can find the field intensity at one mile using an equation given earlier.

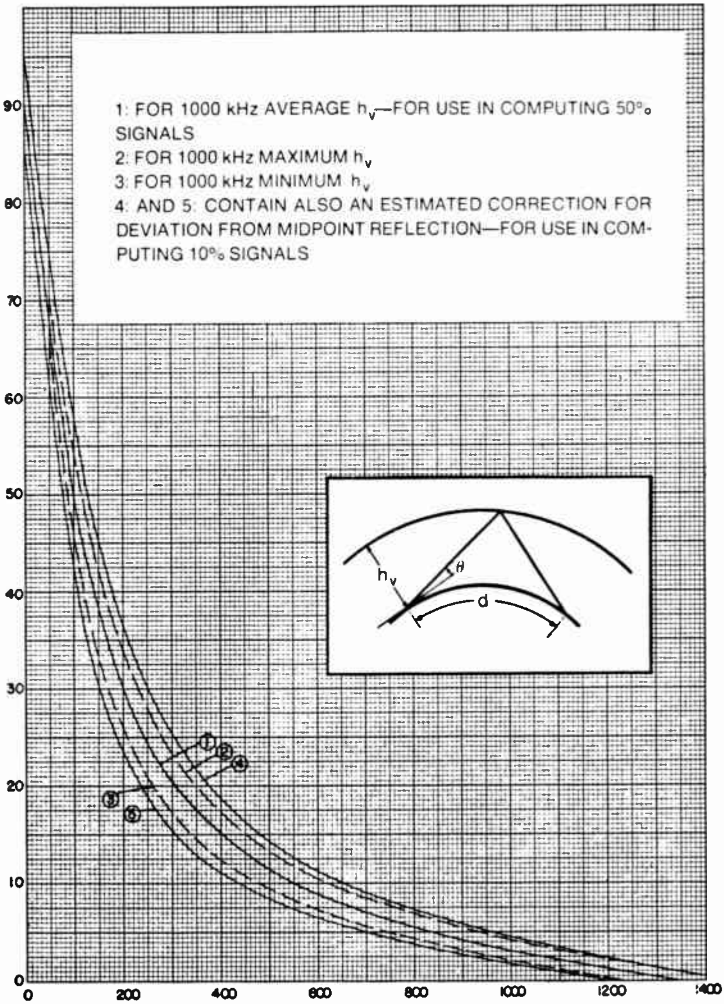


Fig. 6-15. Graph for finding radiation (departure) angle θ as discussed in text.

$$E = \frac{E_0 \sqrt{P}}{d} = \frac{195\sqrt{10}}{1} = 617 \text{ mV/m}$$

where E_0 = effective field intensity of 90° antenna at one mile for 1 kW (Fig. 6-8)
P = actual radiated power
d = distance from the antenna in miles

Now, to find the field intensity at *one mile* at an angle of 11°, we use the contour curves on the right side of Fig. 6-9. From the curve for a 90° antenna, we see that if the field intensity along the surface of the earth was 100 mV/m, the field intensity at an angle of 11° would be about 98 mV/m. This is the same as saying that the radiation at an elevation angle of 11° is 98% of the radiation along the horizon. We just found that in the case at hand the field intensity along the surface of the earth will be 617 mV/m, so the field intensity at an angle of 11° will be

$$617 \times 0.98 = 604 \text{ mV/m}$$

Now we go to the 50% curve of Fig. 6-14, where we see that the 50% field intensity would be about 58 mV/m if the radiated field intensity was 100 mV/m at an elevation angle of 11°. Actually, our field intensity at this angle was found to be 604 mV/m, so the value of field intensity that we would expect to find 50% of the time at a distance of 500 miles would be

$$\frac{604}{100} \times 58 = 350 \text{ mV/m}$$

Finding the probable intensity of skywave signals admittedly isn't a simple procedure, but it is a very handy technique to be familiar with. If an engineer receives a complaint that his station is causing interference to a station in a distant city, the first thing he should do is to find what angle his antenna is radiating at. Then he can start checking to find out what might have gone wrong.

TOP-LOADED AND SECTIONALIZED TOWERS

From the preceding discussions of groundwave and skywave propagation, it is apparent that, for the greatest primary service area with a given radiated power, we should maximize the field intensity along the surface of the earth and minimize the radiation at angles above the horizon. One approach is to use a tower that is nearly 1/2 wavelength high. If the height is increased beyond this, a high-angle lobe will form that can cause skywave interference and reduce the distance to the fading wall.

As noted earlier, the vertical-radiation pattern of an antenna is related to the current distribution along the tower. In Fig. 6-16 we see that when a tower is greater than 180°, there is a phase reversal of the current. This means that current is flowing on the tower in two different directions at the same time. This is responsible for the high-angle lobe that is found in the vertical-radiation pattern in towers that are taller than 180°. If we could find a way to avoid the phase reversal in a tall tower, we could increase the field intensity along the surface of the earth without creating a high-angle lobe in the pattern. This can be done, and many different types of antennas have been designed for the purpose.

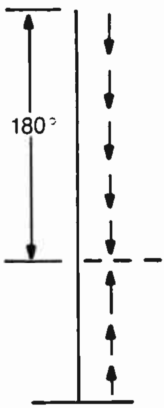


Fig. 6-16. Current direction at one instant.

The simplest antenna of this type is the so-called Franklin antenna, or Franklin array, shown in Fig. 6-17. Here we have two 180° towers, one mounted on top of the other. This arrangement does not act lile a 360° tower, because of the way

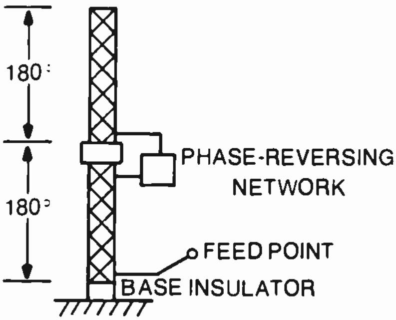


Fig. 6-17. Franklin antenna.

signals are fed to it. The feed system is designed so that the currents in the two 180° sections are in phase. This means that we have no phase reversal in the current distribution and consequently, no high-angle lobe in the vertical-radiation pattern. Several of these antennas have been in use for many years. The principal limitation has been keeping the currents in phase through all the weather conditions that affect the insulation between the two sections. Additional work is being done on sectionalized towers, and they may find wider use in the future.

Another thing that we notice when comparing the radiation pattern with current distribution on antennas of various heights is that, in taller towers, which have greater radiation along the surface of the earth, the *current loop*, point where the current is greatest, is farther from the ground. It would seem, therefore, that if we could find a way to raise the current loop in an antenna, we would also increase the field intensity along the surface of the earth. One way of doing, this is by *top loading* (Fig. 6-18).

In an ordinary tower the current is effectively zero at the top of the tower simply because there is no place for it to flow. Actually, the current is not exactly zero, because there is

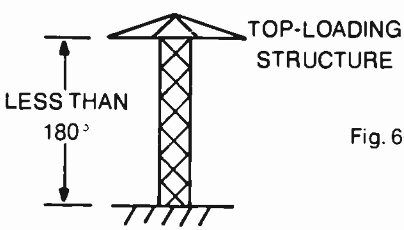


Fig. 6-18. Top-loaded antenna.

always some capacitance between the top of the tower and the earth. This provides a clue as to how top loading works. When there is a structure such as the "top hat" of Fig. 6-18 at the top of the antenna, there will be a substantial current at the top as shown in Fig. 6-19. The top-loaded antenna looks electrically like a taller tower in that the point of maximum current is higher above the ground.

Top-loaded antennas have been used in many cases in the past where, for one reason or another, taller towers were impractical. The combination of top loading and sectionalizing seems to provide an opportunity for improved control of current distribution and hence the vertical-radiation pattern of a tower.

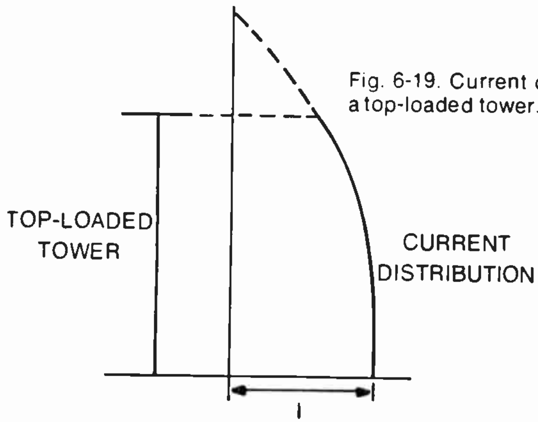


Fig. 6-19. Current distribution on a top-loaded tower.

TOWER IMPEDANCES

Whenever we speak of the impedance of an antenna, we must be careful to specify the point in the antenna that we are talking about. An antenna is in many respects like a transmission line that is open at the receiving end. There are standing waves of voltage and current along the tower; therefore the impedance varies along the tower. At the top of the tower the current is zero, or nearly so, and the voltage is high. The impedance, which is the ratio of voltage to current, is also high at the top. If the current really dropped to zero, the impedance would be infinite. About 90° down from the top of the tower, the voltage is minimum and the current is maximum. The impedance at this point will, therefore, be much lower.

In standard broadcasting we are interested in the impedance at two points on the tower. One is the impedance at the base of the tower, where it is fed. The other is the impedance at the current loop, where the current is maximum. We will consider these two impedances separately.

Base Impedance

We are interested in the impedance at the base of the antenna for many reasons. Inasmuch as the base is where energy is usually fed to a tower, we must match the base impedance to the characteristic impedance of the transmission line for maximum power transfer and minimum reflection. Furthermore, the base impedance includes both the radiation resistance of the tower and the loss resistance assocated with it. If this impedance is extremely low, the loss resistance will be a substantial part of the total resistance, thus making the losses high.

Figure 6-20 shows a plot of the resistance and reactance at the base of a tower versus tower height. In this figure the tower height is given in fractions of a wavelength. The height corresponds to the length of a wave at the operating frequency *in free space*; that is, the velocity of propagation along the tower is not taken into consideration when specifying the height. This might seem like a strange practice. After all, when we specify the length of a transmission line, we take the velocity of propagation into consideration. There are two reasons why we don't do the same thing with antenna height. One is that the velocity of propagation depends on the physical configuration of the tower and often is not known. The other is that the vertical-radiation pattern is easier to compute when the height is specified in terms of the velocity of propagation in free space.

Considering the reactance curve of Fig. 6-20, we see that for antennas that are much shorter than 1/4 wavelength, the reactance is negative, or capacitive. Actually the tower contains both inductance and capacitance at all heights. The reactance is capacitive at tower heights below 1/4 wavelength

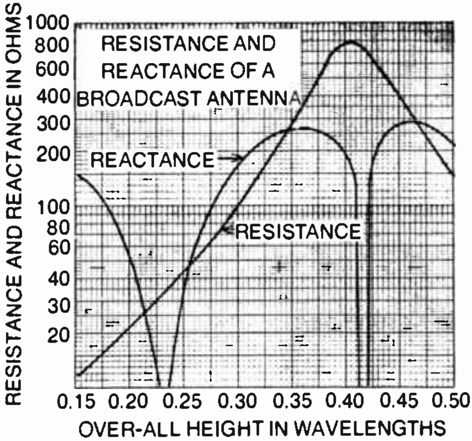


Fig. 6-20. Resistance and reactance versus tower height.

because the capacitive reactance predominates at these heights. As the tower height increases, the capacitive reactance decreases until, at a height just below 1/4 wavelength, the reactance drops to zero. At this height the inductive and capacitive reactances are equal and cancel each other. This height is often called the *first resonance* of the tower. As the tower height is increased above this value, the reactance becomes inductive and increases with tower height. At somewhere below 1/2 wavelength, the reactance again drops to zero. This point is called the *second resonance* of the tower.

Looking now at the curve for the resistive part of the base impedance, we see that it increases with tower height from a small value to a maximum near the second resonance.

Looking at both the resistance and reactance curves of Fig. 6-20, we see that at the first resonance the antenna looks a lot like a series-resonant circuit. The reactance is zero and the resistance is small. At the second resonance the antenna looks like a parallel-resonant circuit. The reactance is zero and the resistance is high. This is very similar to what we would find if we measured the impedance looking into transmission lines of comparable length that were open at the receiving end. Of course, the impedances of an antenna are not exactly the same as those of a transmission line, because an antenna is designed to radiate energy and a transmission line is not.

The ratio of the reactance to the resistance at the base of a tower is of interest because it influences the bandwidth of the tower. We can apply the concept of *Q*—that is, the ratio of reactance to resistance—to antennas as well as circuits. In antenna work we like to keep the *Q* low because high-*Q* circuits and antennas have narrow bandwidths and high losses. In standard broadcast stations we like to keep the *Q* of antennas and networks to not much higher than 3.

Another part of the base impedance that is of interest is the ratio of the resistive component to the characteristic impedance of the transmission line. If the ratio is greater than about 10:1, the design of the impedance-matching network will be complicated somewhat.

From the broadcast engineer's point of view, there isn't much that can be done about the height of a tower. He is interested more in how the impedance of a tower of a given height varies with frequency. Figure 6-21 shows the reactance and resistance seen at the base of a tower that is 90° high at the carrier frequency. Note that the reactance is inductive on either side of the carrier frequency and that the first resonance is just below the operating frequency. The FCC Rules require that the base impedance be measured at 5 kHz intervals over a range of 20 kHz below the carrier frequency to 20 kHz above the carrier frequency.

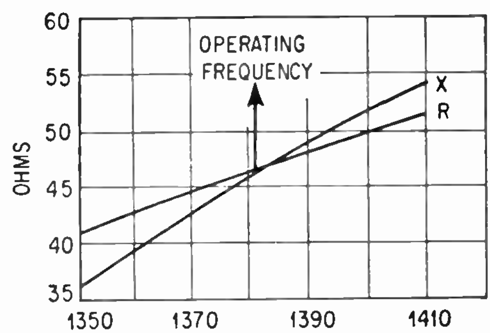


Fig. 6-21. Base resistance and reactance versus frequencv.

Loop Impedance

In directional antennas the magnitude and phase of the current at the current loop, which is about 90° from the top of the tower, are sampled and used in computing the radiation pattern. It is thus useful to have some idea of the impedance of the antenna at the current loop. Unfortunately the relationship between base impedance and loop impedance depends on the actual current distribution along the antenna. This, in turn, depends on many factors, including the shape of the tower and the presence of other structures, such as guy wires.

Figure 6-22 shows several different tower shapes and the current distribution of each. If the tower were infinitely thin, the current distribution would be very nearly sinusoidal. When the cross section of the tower is uniform, the current distribution can still be assumed to be sinusoidal for most practical purposes. When the tower cross section is not uniform, the current tends to increase with cross section.

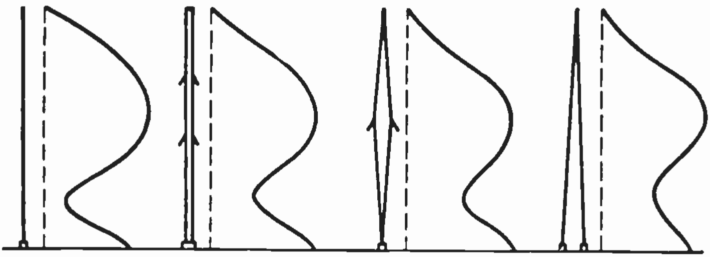


Fig. 6-22. Current distribution for various tower shapes.

When the current distribution is sinusoidal, the base resistance is related to the loop resistance by the equation

$$R_b = \frac{R_l}{\sin^2 h}$$

where R_b = base resistance in ohms
R_L = loop resistance in ohms
h = height of the tower in degrees

This equation neglects reactance, which isn't important, because the base reactance is tuned out by the impedance-matching network.

To a first approximation the base and loop currents are related by the equation

$$Ib = I_l \sin h$$

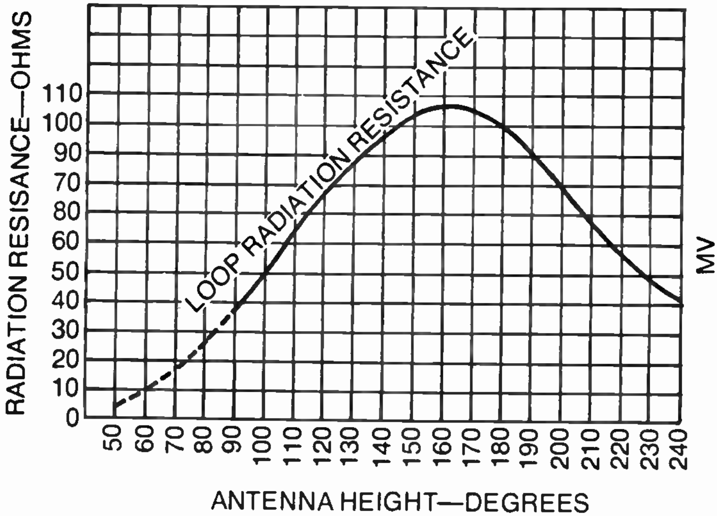


Fig. 6-23. Radiated field and radiation resistance as function of antenna height.

where I_b = base current in amperes
I_l = loop current in amperes
h = tower height in degrees

Figure 6-23 shows a plot of loop resistance for various tower heights. Of course, these curves are only approximate, but they are sufficiently accurate to give the engineer a good idea of the range in which tower impedances should fall.

Chapter 7

Introduction to Directional Antennas

By far the most widely misunderstood antenna system is the AM directional-antenna array. Directional antennas are designed by consulting engineers who specialize in the field, and often all maintenance other than the simplest routine inspection is also done by consulting engineers. The whole field is often thought of by the broadcast engineer as some sort of black art rather than an engineering discipline.

One reason so few broadcast engineers master the principles of directional antennas is that the mathematical expressions involved are usually complicated. Another reason is that many directional-antenna systems seem to resist operating in the way they theoretically should. In this chapter we will develop the equations for determining the field intensity from a directional-antenna system. We will do this by considering each of the parameters that influences the field intensity, one at a time, and then combine all of the parameters into a single equation. In this way the complicated equation, when we get to it, will have lost some of its awesomeness.

There are many parameters that enter into an equation for the field intensity from a directional antenna, including:

1. The geographical orientation of the antenna system
2. The spacing between the towers of the array

3. The bearing from some reference to the point where the field intensity is to be determined
4. The relative phases of the currents in the towers
5. The relative magnitudes of the currents on the towers

Fortunately many of these parameters are fixed when the system is installed and do not have to be varied. For example, there is nothing a broadcast engineer can do about the spacing between the towers of an array once the towers have been erected. What the engineer needs to know is how all of the parameters of a system contribute to the antenna pattern, as well as how to hold the pattern within prescribed limits by varying only the magnitudes and phases of currents in the towers.

Let us start by considering a simple array that has only two towers. To keep things simple, we will assume that the towers are of equal height and are located on a north—south line. We will further assume that the earth in the vicinity of the antenna is flat and that all field intensities we consider are along the surface of the earth. For now we will only be concerned with the shape of the pattern, that is, the field intensity in one direction as compared with that in another direction. In our analysis we will make use of the following two principles, which were described earlier.

Linear Superposition. Inasmuch as our antenna system is a linear system, we can find the field intensity at some point in space by finding the field intensities that each of the towers would produce at that point if acting alone, and then combining the fields to find the resultant field.

Vector Addition. We will consider the field produced by each tower of a directional-antenna array as a vector quantity, that is, a quantity having both a magnitude and a phase angle. We will find the resultant field intensity at a point in space by combining the field intensities from each of the towers by vector addition.

In dealing with the field intensities from the various elements of a directional-antenna system, we will deal with two different sets of values. The first is the theoretical value, which was determined when the array was designed. This is really what the values of parameters ought to be. We will also deal with measured values, which tell us what the values of parameters *actually are*. In this chapter we will only be concerned with theoretical or calculated patterns. In Chapter 14 we will deal with field-intensity measurements and their interpretation.

TOWER-SPACING EFFECTS

Figure 7-1 shows two towers spaced along a north–south line. We will assume for the present that the currents in the two towers are in phase and have the same magnitude. We will represent the field intensity from each tower as $1 \angle 0$, meaning that it has a unit magnitude and that the field intensities are in phase. The tower spacing is represented as the distance S. For our purposes, it will be much more convenient to express S in degrees than in feet. Of course, when we know the operating frequency, we can convert between degrees and feet whenever it is convenient to do so.

Effects in Line with and Perpendicular to Towers

In Fig. 7-1 we have two observation points. Point P_E is due east of the midpoint between the two towers. Point P_W is due west of the midpoint. It is obvious that the signal from tower 1 travels the same distance to P_E as the signal from tower 2 does. Since the fields are in phase at the towers and travel through the same distance, they will be in phase when they reach P_E. Thus the resultant field intensity is the vector sum of the field intensities from the two towers, that is,

$$E = E_1 + E_2$$

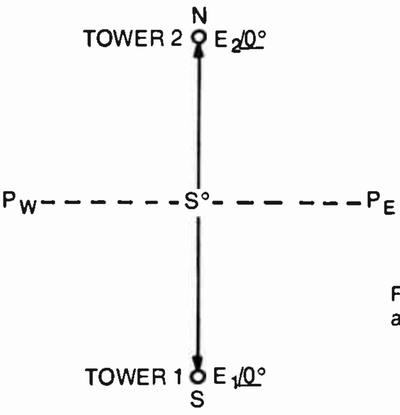


Fig. 7-1. Effect of tower spacing at right angles to line of towers.

(Here we have neglected any attenuation in the paths from the towers to the observation points. We can get around this simplification quite easily later on.)

The paths from the towers to P_E or P_W will still be equal if we move the towers apart or closer together. Thus the resultant field intensity at P_E or P_W is independent of the spacing between the towers. In general, the field intensity along a line perpendicular to the line of towers and passing through their midpoint is independent of the spacing between the towers. That is, we can move the towers closer together or farther apart, but the signal along the midpoint line will not change.

In Fig. 7-2 our observation points are along the line of towers. Point P_N is due north, and point P_S is due south. Point P_N is closer to tower 2 than to tower 1, so the signal from tower 2 arrives at P_N a very short time ahead of the signal from tower 1. This time difference is extremely small, but extremely important. It means that the signal from tower 2 leads the signal from tower 1 by some phase angle. The amount of the phase angle is simply $S°$ where S is the spacing between the towers. It is for this reason that we specify S in degrees instead of feet.

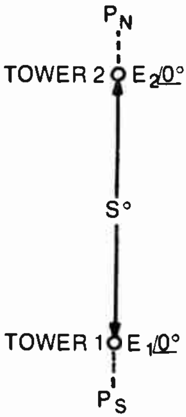


Fig. 7-2. Effect of tower spacing along line of towers.

The field intensity at P_N is the vector sum of the signals from the two towers, but the signals are not in phase, because of the different path lengths. The resultant field intensity at point P_N is

$$E = E_1 \angle 0° + E_2 \angle 0° + S°$$

For example, suppose that the two towers were separated by 180°. The resultant field intensity at point P_N would be

$$E_2 = E_1 + E_2 \angle 180° = E_1 - E_2 = 0$$

Since E_1 and E_2 are equal, there would be no signal at all at point P_N.

A similar situation prevails at point P_S. Here the signal from tower 2 lags the signal from tower 1 by $S°$. If the spacing S was again 180° the resultant field would be

$$E = E_1 \angle 0° + E_2 \angle -180° = E_1 - E_2 = 0$$

Again, there is no signal at all at the observation point.

In both examples above we somewhat arbitrarily assigned a 0° phase angle to the signal from tower 1 and assigned the lead or lag to the signal from tower 2. This merely means that we have chosen tower 1 as the reference tower. It is common practice in directional antenna systems to chose one tower as the reference tower. In most of the examples in this chapter we will designate tower 1 as the reference tower.

Effects at Other Bearings

Before going any further, we must stipulate that our observation points are not close to the towers. We must be far enough away that we are concerned with the radiation field, and not the induction field, from each tower. Next, we must be far enough away that we can consider the array of towers to be a point source of radiation (this is shown in Fig. 7-3). If the observation point *P* is far enough away from the towers, we can consider the lines *A* and *B* to be parallel for all practical purposes. The equations that we will derive for field intensity will all be based on this assumption. This is why directional-antenna field-intensity measurements tend to be meaningless if they are made too close to the antenna site.

So far we have found that tower spacing has no effect at all on field intensity at points east and west of our line of towers, but a very significant effect on the field intensity at points north and south of the towers. The question naturally arises as to what effect the spacing between towers has at observation points at other bearings.

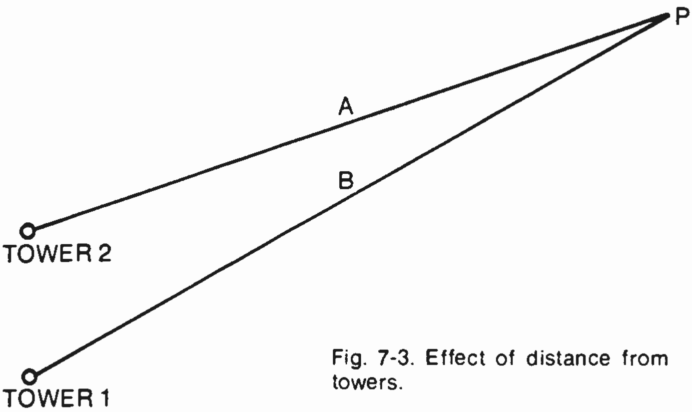


Fig. 7-3. Effect of distance from towers.

In Fig. 7-4 we have the same two towers as in the preceding section, but our observation point is no longer on the line of towers or perpendicular to the line. The observation point P is at an arbitrary angle ϕ to the line of towers. Just by looking at the figure, we can tell that the signal from tower 2 will travel over a shorter path than that from tower 1, so it will lead the signal from tower 1, in phase. We want to find out just how much phase difference there is between the two signals, and particularly how this phase difference is influenced by the spacing between the two towers. We have already assumed that the observation point is far enough away from the towers that we can consider the line from tower 1 to point P to be parallel to the line from tower 2 to point P. This assumption allows us to draw the right triangle shown in Fig. 7-4. From this figure we see that the path difference in electrical degrees between the signals is given by

$$\text{Path difference} = S \cos \phi$$

where S is the spacing in electrical degrees, and ϕ is the bearing from the line of towers in angular degrees. In this expression we use both angular and electrical degrees to find a phase shift in electrical degrees. This should cause no confusion. In fact, the choice of these units simplifies the calculations.

The expression $S \cos \phi$ will give the relative field intensity at any angle, even on the north–south and east–west lines that we investigated earlier. If the observation point is on an east–west line, the angle ϕ will be $\pm 90°$. The cosine of 90° is zero, so the equation tells us that the spacing between towers has no effect on the signal at points due east and west of the towers. This is exactly what we saw in our inspection of Fig. 7-1. If the observation point is due north or south of the towers, the angle ϕ is 0° or 180°. At these values of ϕ, $\cos \phi$ is either +1 or −1. This means that the phase difference between the signals from tower 1 and tower 2 is either $+S$ or $-S$, which is exactly what we found in our inspection of Fig. 7-2. Thus the expression $S \cos \phi$ is general and can be used to find the relative field intensity at any bearing from the line of towers.

This expression gives some additional insight into the pattern of a 2-tower array. It shows that the relative distance traveled by signals from the two towers involves the function $\cos \phi$. An angle may have either a positive or a negative value without changing the value of its cosine. That is, $\cos 45° = \cos -45°$. This means that the pattern on one side of the line of towers will be exactly the same as the pattern on the other side; the pattern will always be symmetrical about the line of towers. This is true of any directional-antenna array where all of the towers are in a straight line, regardless of the number of towers. If we have a null on one side of the line of towers, we will also have a null on the other side at the same angle. Of course, the engineer operating a directional antenna can't do anything about the direction of the line of towers once the system is installed. The designer, however, can and does take advantage of the symmetry to obtain the desired pattern.

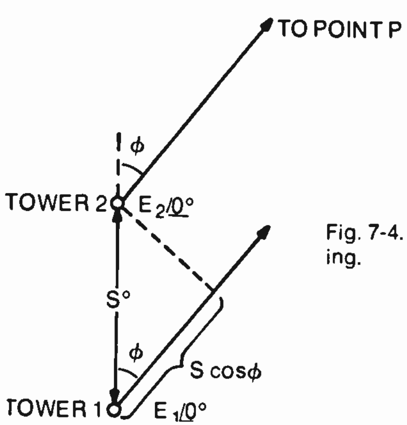


Fig. 7-4. Effect of spacing at bearing.

Effect of Phase of Tower Currents

Figure 7-5 shows a 2-tower array in which we will assume that the currents in the two towers are equal in magnitude but not in phase. Tower 1 is to be our reference tower, so the phase angle of its current is 0°. The phase of the current in tower 2 will be $\gamma°$ earlier or later than at tower 1. If the sign of γ is +, the signal arrives at tower 2 earlier than at tower 1. If the sign of γ is −, the signal arrives at tower 2 later than at tower 1.

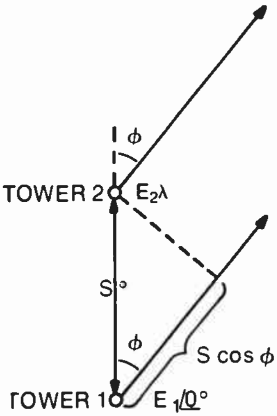


Fig. 7-5. Effect of phase difference between tower currents.

Note that the phase angle γ is not dependent on any of the other parameters in the system. By means of the phasor in the feeder system, we can make γ assume any positive or negative angle that we wish. The effect of this phase shift is to change the relative time required for the signals from towers 1 and 2 to reach any observation point. One thing that this additional parameter does for us is to give us an infinite variety of patterns.

To see this a little better, let's put some numbers into our reasoning. Let's assume that all of our observation points are at the same radial distance from the center of our line of towers and that they are far enough away that we can consider the array a point source. The signal from tower 1 at any observation point will be

$$E = E_{\cdot} \angle 0°$$

The field intensity from tower 2 is given by

$$E = E_2 \angle S \cos \theta + \gamma$$

The phase angle of $S \cos \phi + \gamma$ includes the effect of the tower spacing S, the bearing angle ϕ, and the phase difference γ between the currents in the towers, all expressed in degrees. We can now write an equation for the field intensity at any point, the only limitation being that the points all be at the same radial distance from the center of the array, and on the surface of the earth.

$$E = E_1 \angle 0° + E_2 \angle S \cos \phi + \gamma$$

One very interesting property of the preceding equation is that the field intensity at an observation point will be zero whenever

$$S \cos \phi + \gamma = \pm 180°$$

This means that for almost any value of γ, we will have two nulls, one on either side of our line of towers. (At some values of S and γ there will be one null off one end of the line of towers.) In Fig. 7-6 we have the pattern of a 2-tower array where the spacing S between the tower is 120°, and the phase angle γ between the tower currents is 100°. The pattern is a plot of the data given in the figure. Note that because the radiation pattern is symmetrical about the line of towers, it is only necessary to compute the pattern for angles between 0° and 180°.

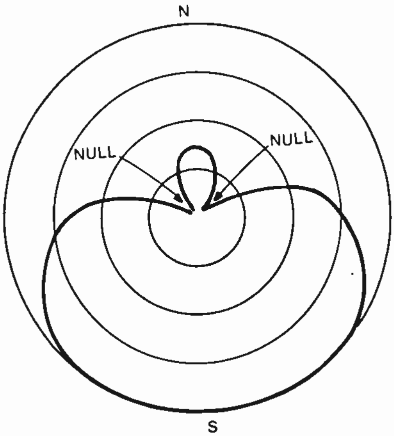


Fig. 7-6. Pattern of 2-tower array.

At this point we can note some other things about the patterns that we can get from a 2-tower array. If the quantity $S \cos \phi + \gamma$ is large enough, we will get two nulls, which can be moved by changing the phase of the currents between the two towers. If the quantity is made much larger, as when the towers are very widely spaced, there will be more than two nulls.

In the data from which we plotted Fig. 7-6 (see Table 7-1), the field intensity at all points has a phase angle. This is not important as far as the pattern is concerned, because the receiver has no way of "knowing" what the phase angle is.

There is a case where we are interested in the phase angle of the field intensity from a 2-tower array, and that is when we wish to combine the patterns from two such arrays to form a more complicated pattern.

0	120 cos 0 + 100	1 Lo + 1120 cos 0 + 100
0	220	0 7/−7°
10	218	0 7/−71°
20	213	0 6/−73 5°
30	204	0 4/−78°
40	192	0 2/−84°
45	185	0 1/−87 5°
47	182	0 03/−89°
48	180	0
49	179	0 02/89 5°
50	177	0 05/88 5°
60	160	0 35/80°
70	141	0 67/70 5°
80	103	1 25/51.5°
90	100	1 29/50°
100	79	1 54/39 5°
110	59	1 7/29 5°
120	40	1 9/20°
130	23	2 0/11.5°
140	8	2 0/4 0°
150	−4	2 0/−2 0°
160	−13	2 0/−6 5°
170	−18	2 0/−9 0°
180	−20	2 0/−10°

Table 7-1. Data for Plotting the Pattern of Fig. 7-6.

FILLING IN NULLS

The pattern in Fig. 7-6 has nulls at bearings of 48° and 312°. In these nulls the signals from the two towers of the array completely cancel each other; theoretically, there is no signal at all along these bearings. In a practical directional antenna, we do not have these complete nulls. There are several reasons for this. In the first place, an array with a complete null would be very difficult to adjust and very unstable. Second, although we may wish to reduce the field intensity along certain bearings to protect the service areas of distant cochannel or adjacent-channel stations, we rarely need to suppress the signal completely. A substantial portion of the population in the primary service area of our station may live in the direction of the nulls. In this case, we want to fill in the nulls enough to provide service to these people. Finally, a perfect null is usually impossible to obtain, because reradiation from other objects, such as guy wires and other structures, prevents complete cancellation of the signals.

To see how we might fill in a null a little, let's look at how we got the nulls in the preceding example. The reason the field intensities canceled completely along certain bearings is that the signals were equal in magnitude and 180° out of phase. If we were to keep the phase difference between the two signals 180° but make their magnitudes unequal, the signals would still tend to cancel, but the cancellation wouldn't be complete. We would have a low field intensity in the direction where formerly there was complete cancellation of the signal. Obviously, the greater the difference in the signals from the two antennas, the greater the field intensity in the null.

Strictly speaking, the term *null* refers only to those bearings where there is complete signal cancellation. We should speak of a *minimum* instead of a *partially filled null*. However we will continue to adhere to the broadcasting custom of calling pattern minima *nulls*.

Field Ratio and Minimum Depth

To specify how much a null will be filled by making the currents in the two towers unequal, we will make use of the concept of the *field ratio* F_{21}. This is simply the ratio of the field intensity from tower 2 to the field intensity from tower 1 or

$$F_{21} = \frac{E_2}{E_1}$$

When the two towers have the same height, the field ratio is equal to the ratio of the currents in the two towers. This ratio is usually designated as M_{21} and is given by

$$M_{21} = \frac{I_2}{I_1}$$

We can write the equation for the field intensity from the two towers as

$$E = E_1 \angle 0^\circ + E_2 \angle \beta$$

where β is the difference in phase between the fields from the two towers. We can define the *relative field intensity* at any point as the ratio of the field intensity produced by an array to the field intensity that would be produced by tower 1 acting alone with the same value of radiated power. That is

$$\frac{E}{E_1} = 1\angle 0^\circ + \frac{E_2}{E_1}\angle\beta$$

Since $F_{21} = E_2/E_1$, we can write this equation as

$$\frac{E}{E_1} = 1\angle 0^\circ + F_{21}\angle\beta$$

By a rather lengthy manipulation, we can rewrite this equation as

$$E = E_1\sqrt{2F_{21}}\sqrt{\frac{1+F_{21}^2}{2F_{21}} + \cos\beta}$$

The term $(1 + F_{21}^2)/(2F_{21})$ is called the *minimum-depth* term. It is a measure of how much the nulls will be filled in. It is interesting that the term will not change value if we replace F_{21} with $1/F_{21}$. For example, we can use 2 or 0.5 as the value of the field ratio and get 1.25 as the value of the minimum-depth term in either case. That is,

$$\frac{1+2^2}{2(2)} = \frac{1+0.5^2}{2(0.5)} = 1.25$$

The minimum-depth term is always equal to 1 or more. When it is equal to 1, the null is perfect. At other values the null is filled in accordingly. Since in the expression for the minimum-depth term it makes no difference whether we use F_{21} or $1/F_{21}$, when we fill a null we need only concern ourselves with the *ratio* between the two tower currents. It makes no difference which tower has the larger current. For this reason the FCC Rules governing directional-antenna systems place a limit on currrent ratios rather than current magnitudes.

COMPLETE PATTERN SHAPE

Now let's see if we can put all that we have discussed so far together to find the pattern of a 2-tower directional-antenna array. Before going further, we should note that in FCC documents the reference point for a pattern is always the center of the array, not necessarily one of the towers. Thus the origin of a pattern is a point midway between the two towers in a 2-tower array, the middle tower of a symmetrical 3-tower array, or the center of a parallelogram array.

The pattern is plotted on polar-coordinate paper, with the center of the paper representing the center of the array. We can rewrite the equation for the field intensity in a form that places the reference point midway between the towers (Fig. 7-7).

$$E = E_1 \angle \frac{\beta}{2}\cos\phi + \frac{4}{2} + E_2 \angle -\frac{\beta}{2}\cos\phi - \frac{4}{2}$$

Inasmuch as our pattern is symmetrical, we only have to calculate the field intensities on one side of the line of towers.

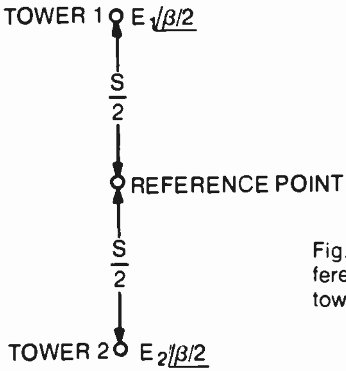


Fig. 7-7. Two-tower array with reference point midway between towers.

True Bearing of a Pattern

So far we have considered that our towers were along a north–south line. In practice, the line of towers is often at some other angle (Fig. 7-8). Of course, this needn't cause any confusion, because orienting the line of towers along any other bearing is just the same as picking up the pattern in Fig. 7-6 and rotating it. When we are calculating the locations of nulls, or the field intensity at any bearing, it is convenient to use the line of towers as a reference. By simply adding or subtracting the angular orientation of the line of towers with respect to north, we can get the geographic location of the various features of the pattern.

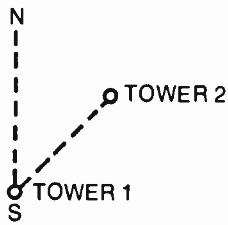


Fig. 7-8. Change of reference axis.

When a bearing is specified with respect to north, it is usually called an *azimuth*, or *true bearing*. Thus a radial line 20° from the line of towers might be referred to as a *bearing of 20 degrees*. And a bearing of 20° with respect to true north, rather than with respect to the line of towers, would be called *20° azimuth* or *20° true*.

The nulls of a pattern should be converted to true bearings to ensure that the primary service areas of other stations are properly protected. The geographic location of other stations will be known in terms of the bearing from true north, not in terms of the bearing of the line of towers.

Putting Numbers on the Pattern

We have developed the general shape of the pattern of a 2-tower array and have seen how the tower spacing and relative magnitude and phase of the tower currents affect the pattern. The patterns we worked with however were relative rather than absolute. They showed the general shape of the pattern but gave us no idea of the actual field intensity in millivolts per meter at any bearing. We will now put numbers on the patterns so that they will actually be plots of field intensity.

Our equation for the effective field intensity at one mile can be written

$$E = E_1 \, [f_1 \, (\theta) \angle \beta_1 \; + F_{21} \, f_2 \, (\theta) \angle \beta_2 \,]$$

where E_1 = effective field intensity from tower 1 at one mile from antenna $f_1(\theta)$,

$f_2(\theta)$ = vertical-radiation characteristic of towers 1 and 2.

F_{21} = field ratio

β_1, β_2 = phases of fields from towers

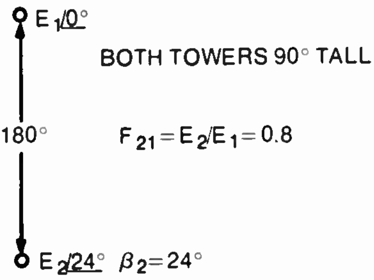
In this equation we have introduced two new terms, $f_1(\theta)$ and $f_2(\theta)$. These represent the vertical-radiation characteristics of our two towers, which were discussed in Chapter 6. For now, we will consider each of these terms to be equal to 1, because we are now only interested in the radiation along the surface of the earth. Thus the only problem remaining in specifying the field intensity at any bearing is to find the inverse or unattentuated field intensity at one mile from each tower. We will still consider the two towers to be identical: so if we find this field intensity for one tower, it will apply to the other tower as well.

We can find the unattenuated field intensity at one mile from a tower in a directional-antenna system as we found it for a single vertical tower in Chapter 6. Using the curve in Fig. 6-8, we can find the field intensity for 1 kW of radiated power. We can then use that figure to determine the actual field intensity for whatever power is being radiated.

Suppose, for example, that we have the 2-tower array of Fig. 7-8 and that both towers have a height of 90°. The total radiated power is 5 kW. From Fig. 6-8 we find that the unattenuated field intensity at one mile from the antenna is 195 mV/m for a radiated power of 1 kW. The field intensity is proportional to the square root of the radiated power, so the unattenuated field at one mile for a radiated power of 5 kW is $195\sqrt{5} = 436$ mV/m. The actual field intensity is then

$$E = 436[1 \angle 0° + F_{21} \;\angle S \cos \phi + \gamma$$

Inasmuch as the two towers have the same height, the field ratio F_{21} will be the same as the ratio of the currents in



$$E = 436[1\underline{/0^\circ} + F_{21}\underline{/18\cos\phi + 24^\circ}$$

ϕ	$\underline{/180\cos\phi + 24^\circ}$	$1\underline{/0^\circ} + 0.8\underline{/180\cos 24^\circ}$	E
10	201	0.39	168
20	193	0.29	124
30	179	0.20	87
40	162	0.34	150
50	140	0.65	283
60	114	0.99	434
70	86	1.33	579
80	55	1.60	696
90	24	1.76	768
100	−7	1.80	783
120	−66	1.51	660
130	−92	1.26	550
140	−114	0.99	434
150	−132	0.76	330
160	−145	0.57	249
170	−153	0.46	200
180	−156	0.42	184

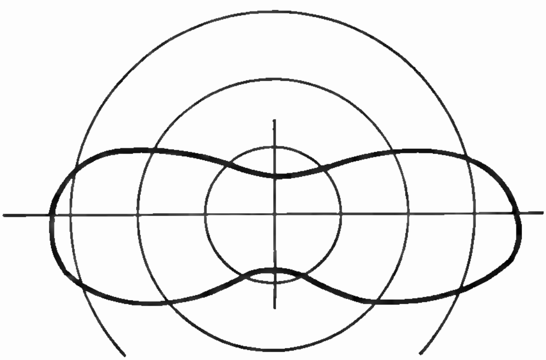


Fig. 7-9. Two-tower array with pattern and data.

the two towers—in this example, 0.8. We can now solve the equation for various bearings and get the field intensity at each bearing. This will be the unattenuated or effective field intensity at each bearing at a radial distance of one mile from the center of the array. The data and pattern are given in Fig. 7-9. The equation for the pattern calls for two partial nulls. The result of the unequal currents in the tower is that the nulls are filled to such an extent that the pattern varies smoothly instead of having sharp notches.

PATTERN SIZE

The term *pattern shape*, as used in the preceding discussion, is self-explanatory. The term *pattern size* is apt to be confusing. It is a measure of how much power is radiated by the antenna system. There are several ways that we can specify the size of a pattern. One way is to specify the *root-mean-square* or *rms* value of the field intensity. The rms value of an antenna pattern is equal to the radius of a circle plotted to the same scale that has an area equal to that enclosed by the pattern. Figure 7-10 shows the pattern of an array with 90° spacing and 90° phasing. The rms value is represented by the radius of the dashed circle. The units of the rms value are the same as the units of the pattern, usually millivolts per meter.

There are several other ways that we can find the rms value of a pattern. One is to use a *planimeter*, which is a drafting instrument that, when moved along the contour of a closed curve, indicates the area enclosed by a closed curve. If

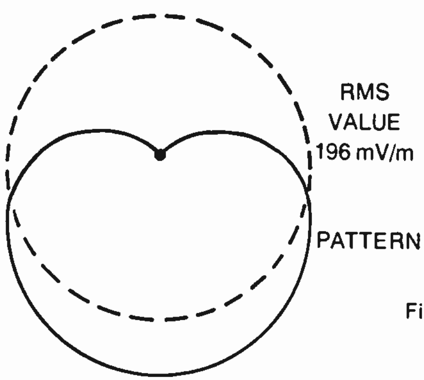


Fig. 7-10. Pattern with rms value

such an instrument isn't available, the area can be found by plotting the pattern on rectangular graph paper and counting the squares enclosed by the pattern. When this is done, the scale of the rectangular graph paper should be the same as the radial scale of the polar graph paper on which the pattern was originally plotted.

The FCC Rules provide that the rms value of the pattern have a certain minimum value, as given in Chapter 6. Using the curves in Chapter 6, we found that a 90° tower radiating 1 kW of power will produce a field intensity of 195 mV/m at a radius of one mile from the tower. If two 90° towers in directional array radiate a total power of 1 kW, it might seem that the rms value of their pattern should also be 195 mV/m. This assumption, in general, is not correct, because the vertical-radiation characteristic of two towers in an array is not the same as that of either of the towers acting alone. An array may "squeeze down" the pattern to increase the radiation along the surface of the earth. By the same token, it may distort the pattern so that the radiation along the surface of the earth is less than would be radiated by one of the towers radiating the same power. This shows the reason for finding the rms value of a directional-antenna pattern.

The rms value of a pattern depends on tower height, spacing, and phasing. Figure 7-11 shows the rms field intensity for a 2-tower array with 90° towers for various values of spacing and phasing. This shows specifically how the choice of

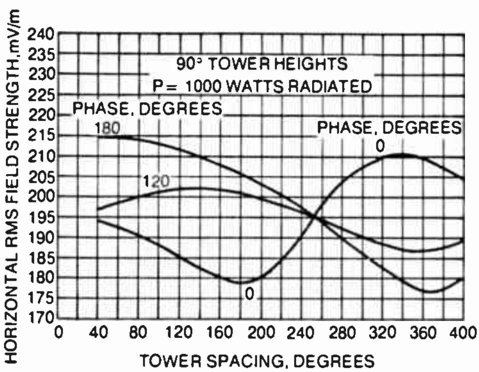


Fig. 7-11. Horizontal RMS field intensity as a function of tower spacing and phasing.

spacing and phasing may produce an array that actually has an rms gain over a single tower. In general, this is a design consideration rather than an operational one. The broadcast engineer should, however, be aware of the consideration.

VERTICAL-RADIATION PATTERN OF A DIRECTIONAL ANTENNA

The vertical-radiation pattern of a directional antenna is of particular importance to stations that operate at night. The skywave signal from such stations may easily cause interference in the primary service area of cochannel or adjacent-channel stations. There are two factors that enter into the vertical-radiation pattern of a directional antenna. The first is the vertical-radiation pattern of each of the towers in the array (Chapter 6). The second factor is the spacing of the towers in the array, which affects the vertical radiation similarly to the way it affects the surface pattern.

Figure 7-12 is a side view of a 2-tower array. The actual spacing between the two towers is $S°$. Now let us look down on the array from point P, which is off the picture at the upper-right side. As with our other observation points, point P is far enough away that the lines from the two towers can be considered to be parallel. The spacing between the towers as seen from point P is no longer S, but appears to be shortened to $S \cos \theta$, where θ is the elevation angle in degrees.

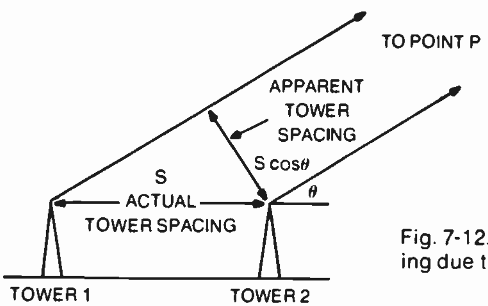


Fig. 7-12. Foreshortening of spacing due to elevation angle.

Now we can write a very general equation for the pattern of a 2-tower directional antenna that applies not only on the surface of the earth but at any value of elevation angle θ. In vector form the equation is

$$E = E_1 \ f_1(\theta) \ \underline{/ S_1 \ \cos\theta \cos\phi + \gamma_1} + E_2 \ f_2(\theta) \underline{/ S_2 \ \cos\theta \cos\phi + \gamma_2}$$

We are now in a position to find the elevation angle θ at which a null occurs. This is important because many directional antennas are designed to place a null in the vertical-radiation pattern to protect the primary service area of some other station. The field intensity above the surface of the earth is given by

$$E = E_1 \; f_1 \,(\theta) \; \angle 0° + E_2 \; f_2 \,(\theta) \quad \underline{/S_2 \; \cos\phi \cos\theta + \gamma_2}$$

We know that the field intensity will never be zero if E_1 and E_2 are not equal. Its minimum value is given by the minimum-depth term

$$\frac{1 + F_{21}{}^2}{2F_{21}}$$

Its bearings are

$$S_2 \; \cos\phi \cos\theta + \gamma_2 \;\; = \pm 180°$$

Thus, with a tower spacing S of 90° and a tower phasing γ of 100°, the pattern will have a minimum value when

$$90 \cos\phi \cos\theta + 100° = \pm 180°$$

This equation tells us several things. First, the null doesn't occur at the same bearing as we look at the antenna from different elevation angles. On the surface of the earth the elevation angle θ is zero, so $\cos\theta = +1$. Thus the null occurs at the angle where

$$90 \cos\phi = 180 - 100 = 80°$$
$$\cos\phi = 90/80 = 0.89$$

This is at a bearing of 27°. As elevation increases, the null rotates and occurs at different bearings. For example, at an elevation of 20°, the null occurs at a bearing (angle to the antenna) of about 19°.

From the foregoing we can see that even though we measure a null on the surface of the earth in the direction of a cochannel station, it is entirely possible to be nearly blasting the station off the air with a skywave signal.

TOWERS OF UNEQUAL HEIGHT

In all of our computations we have used towers of equal height. This not only simplified our computations, but it is typical of most directional-antenna systems. Under some circumstances, however, the towers in an array are not all the same height. This often occurs when an FM or TV antenna is mounted on top of one of the towers after the system has been installed. The chief effect of unequal-height towers is that the field ratios from the two towers will no longer be equal to the current ratios. Suppose, for example, that we have one tower 90° in height and another 112° in height. Referring to Fig. 6-8, we see that for 1 kW of power a 90° tower will provide an effective field intensity at one mile of 195 mV/m, and a 112° tower produces a field intensity of 202 mV/m. To get equal field intensities, we must reduce the current in the 112° tower. The field intensity is directly proportional to the current in the tower; therefore, the current I_2 in the 112° tower must be reduced by multiplying by

$$\frac{195}{202} = 0.97$$

Suppose that with these two towers we wish to produce a field ratio of 0.8. We can no longer make the current ratio 0.8, but must modify it by the figure that we just computed. Thus the current ratios must be

$$\frac{I_2}{I_1} = 0.97 \times 0.8 = 0.78 \qquad \frac{I_1}{I_2} = 1.21$$

We can use either figure because, as far as the pattern is concerned, it doesn't matter which tower has the larger current.

THEORETICAL AND STANDARD

Several types of patterns are plotted for directional antennas. Some of these are theoretical, and others are *empirical*, that is, based on actual measurements rather than computations. Today the theoretical patterns are often actually determined on a digital computer by the designer. This saves a tremendous amount of labor in checking proposed designs to see if they will provide adequate protection to other stations and adequate coverage of the primary service area.

The broadcast engineer will not be concerned with preparing theoretical patterns unless he is making a major modification to this system. He should, however, be conversant with how they are prepared and what they mean.

The FCC Rules require a plot of the theoretical pattern of the signal strength along the surface of the earth. This is a plot of the unattenuated or effective field intensity at a distance of one mile from the center of the array. When vertical radiation is significant as when a station is on the air at night, a similar pattern must be plotted for elevation angles up to 60°, with a separate pattern for each increment of 5°.

Patterns must be plotted on polar graph paper of standard letterhead size. The graph area is then 7 by 10 in. The pattern must be oriented with 0° corresponding to true north, and not to the line of towers or any other reference. The scale divisions should be 1, 2, 2.5, or 5. Any field intensity on the pattern that is less than 10% of the effective field must be plotted on an expanded scale. A typical pattern is shown in Fig. 7-13. Note that the low field intensities to the south of the antenna are plotted on a × 10 expanded scale.

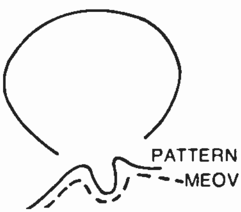


Fig. 7-13. Pattern with MEOV (maximum expected operating values).

Prior to January 18, 1971, the theoretical pattern was computed using some variation of the equations presented earlier in this chapter. The resultant pattern was similar to that of Fig. 7-13. After the pattern was plotted, the engineer made allowances for the typical deviation of directional antennas from theoretical performance. These allowances were made by increasing the pattern by roughly 5% in critical nulls. The new values were then added in the pattern as shown by the dashed line in Fig. 7-13. The values represented by the dashed line are called MEOV, or *maximum expected operating values*. MEOVs were used by the FCC in determining the amount of protection provided to other stations. Some engineers, just to be safe, would sketch an MEOV curve around the entire pattern.

If, during the proof of performance of the antenna system, the signals in the nulls could not be brought to the theoretical values, they might still be within the MEOV and thus meet the performance standards set forth in the station's construction permit. Thus we actually had two patterns—the theroetical pattern and the pattern represented by the MEOV curve.

As of January 18, 1971, the FCC decreed that every station should have one pattern, the *standard pattern*. The standard pattern is actually the old theoretical pattern, modified by the addition of two factors. One factor is 2.5% of the rms of the fields of the individual towers, or 6.0 mV/m, which ever is greater. This factor is added to the theoretical value by what is called *quadrature addition*. This means that both factors are squared and then added together, then the square root is taken. The resultant figure is then multiplied by 1.05 to get the standard radiation pattern.

The equation given in the FCC Rules for the standard pattern is

$$E_{std} = 1.05\sqrt{E_{th}^{2} + Q^{2}}$$

The terms of this equation are discussed below.

The field intensity E_{std} of the standard pattern represents the expected unattenuated field at one mile.

The theoretical pattern E_{th} is calculated with an assumed loss resistance of one ohm at the current loop of any tower over 90° high, or at the base of an antenna less than 90° high.

The quantity Q is the greater of the following:

$$0.025\, f(\theta)\, E_{rss} \text{ or } 6.0\, f(\theta) \sqrt{P_{kW}}$$

The vertical form factor $f(\theta)$ is for the shortest tower in the array. This figure is taken from Fig. 6-9 and is used for plotting patterns at vertical angles. For the horizontal pattern it is 1.

The symbol E_{rss} stands for the *root-sum-square* or *rss* value of the field intensities of the towers of the array. The rss value should not be confused with the rms value. The rss value of a number of quantities is the square root of the sum of their squares (the absolute value of an impedance is found from resistance and reactance).

The power input to the array expressed in kilowatts is represented by P_{kW}. If the power is less than 1 kW, the quantity 1 is used here.

The FCC also requires that the rms value of the pattern be given. This is computed from the theoretical pattern in the above equation.

The standard radiation pattern was adopted to do away with MEOV, which were inconvenient. The goal was for every station using a directional antenna to have only one pattern. Allocations and interference contours would then be based on patterns that were all calculated in the same way. The Commission has, however, recognized that directional-antenna design is fraught with difficulties that cannot always be anticipated. If, when actual measurements are made of the radiation pattern, it is found that the radiation exceeds the standard pattern over a limited range, provision is made for *augmenting* the pattern. Augmentation is somewhat similar to adding an MEOV curve, except that a definite procedure is given for computing the augmentation, and the augmented pattern replaces the original standard radiation pattern. Hence there is still only one pattern for each station, which, in some instances, may be an *augmented* pattern.

WRAP UP

For the past several pages we have investigated the properties of the 2-tower directional-antenna array. Our purpose has not been to achieve design capability, but rather to acquire an understanding of how simple arrays operate. And much of the information that we developed in connection with the 2-tower array can be applied directly to more complex arrays.

Chapter 8

Complex Directional—Antenna Arrays

The 2-tower directional-antenna array described in the preceding chapter is capable of producing a wide variety of radiation patterns and meets the requirements of many standard broadcast stations. Quite a few other stations, however, require an antenna pattern that protects the service areas of many other stations while providing adequate coverage of their own service areas. This type of pattern is obtained by using more than two towers. Nine or more towers may be used in an array to obtain nulls at many different angles or to form very broad nulls in a particular part of a pattern.

Regardless of the number of towers used in an array, the field intensity at any point in space may be found by first finding the field intensity that would be produced by each tower alone, then taking the vector sum of the several field intensities. Unfortunately, as the number of towers in an array increases, the complexity of the field equations also increases at a disturbing rate. The equations are not necessarily difficult to comprehend, but due to the large number of terms, solution is often tedious.

Much of the complexity in the mathematical work is in getting the equations into a form for solving with a slide rule, pencil, and paper. With an electronic calculator vector addition is less tedious. We will leave the terms of the equations that we use in vector form because it keeps them much simpler and makes them easier to solve with an electronic calculator.

In this chapter we will develop a method of finding the field intensity at any point in space from an array containing any number of towers. We will consider a graphical technique that will show how the field from each tower contributes to the field intensity at any point in space. We will also consider examples of patterns that may be obtained from 3- and 4-tower arrays.

SPACE REFERENCES

In a 2-tower array the towers obviously lie in a straight line. When more than two towers are used, they may or may not all be in line. For this reason we must develop a reference system that will let the towers of an array be located anywhere with respect to each other. We do this by assigning a space-reference point for the array, which isn't necessarily located at any of the towers. We also assign a space-reference axis, which is a north–south line through the space-reference point. This is shown in Fig. 8-1.

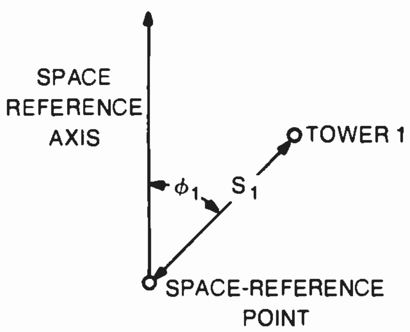


Fig. 8-1. Space-reference system.

With this system there is a spacing S and an azimuth angle ϕ associated with every tower in the system. In the figure, S_1 is the spacing between tower 1 and the space-reference point.

The angle ϕ_1 is the angle between the space-reference axis and a line from the space-reference point to tower 1. If one of the towers of an array was located at the space-reference point, its spacing and azimuth angle would be zero. This is often true of the center tower of a 3-tower array.

FINDING THE FIELD INTENSITY OF A THREE-TOWER ARRAY

To intelligently operate and maintain a directional-antenna system, the engineer must know the relationship between the parameters of the system and the radiation pattern. The parameters over which he has control are the amplitude and phase of the currents in the various towers. He is most interested in the field intensity at the monitoring points set forth in the station license. In general, the magnitudes and phases of all of the currents in all of the towers influence the field intensity at all of the monitoring points. The amount of influence that each current amplitude and phase has on the field intensity at each of the monitoring points depends on all of the design parameters of the array, including the tower spacing and orientation.

We will start our analysis with the 3-tower in line array shown in Fig. 8-2. The towers are conveniently located on the space-reference axis. The center tower is located at the space-reference point of the array and is designated *No. 1.* We wish to find the effective field intensity at point *P*, which is located at a distance of one mile from the space-reference point, at an angle of 40° from the space-reference axis. We know from the superposition principle that the field *E* at point P will be given by an equation of the form

$$E = E_1 \angle \beta_1 + E_2 \angle \beta_2 + E_3 \angle \beta_3$$

where E_1, E_2, E_3 = field intensities of towers acting alone
$\beta_1, \beta_2, \beta_3$ = phase angles of fields from towers, at point *P*

Since we can easily solve an equation of this type with an electronic calculator without further manipulation, it is only necessary for us to find values for the *E*s and βs to be able to find the effective field intensity at point *P*. The *E*s are fairly

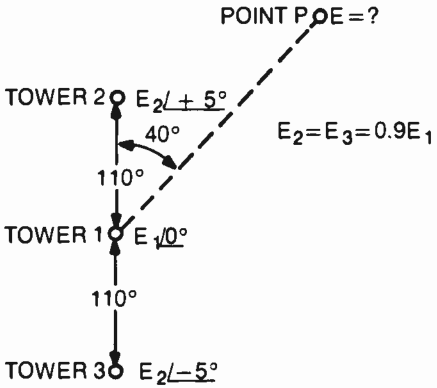


Fig. 8-2. Three-tower array.

easy to find. From the information on the station license and from meter indications, we can compute the effective field intensity that each tower would produce from the techniques given in Chapter 6.

Effective Field of Each Tower

Suppose, for example, that all of the towers are 90° in height and that the total radiated power is 5 kW. Inasmuch as the towers are all the same height, the field ratios will be equal to the current ratios, which are specified on the license and can be measured with the antenna monitor. We will assume that the ratios are as shown in Fig. 8-2, that is,

$$E_2 = E_3 = 0.9\,E_1$$
$$I_2 = I_3 = 0.9\,I_1$$

The power radiated by each tower is proportional to the square of its current. Therefore we can compute the power from each tower as follows.

$$P_1 + 0.9^2\,P_1 + 0.9^2\,P_1 = 5\,\text{kW}$$
$$P_1 = 5/2.62 = 1.91\,\text{kW}$$
$$P_2 = 0.81\,P_1 = 1.55\,\text{kW}$$
$$P_3 = 0.81\,P_1 = \underline{1.55\,\text{kW}}$$
$$5.0\,\text{kW Total power}$$

Knowing the power radiated by each tower, we can use the information in Fig. 6-8 of Chatper 6 to find the effective field at one mile for each of the towers. From Fig. 6-8 we see that if the radiated power was 1 kW, the effective field at one mile from each of the towers would be about 195 mV/m. Inasmuch as the field intensity is proportional to the square root of the power, we can now find the effective field for the actual radiated power.

$$E_1 = 195\sqrt{1.91} = 271 \text{ mV/m}$$
$$E_2 = 195\sqrt{1.55} = 244 \text{ mV/m}$$
$$E_3\ 195\sqrt{1.55} = 244 \text{ mV/m}$$

This gives numbers that we can substitute for the Es in our equation, which becomes

$$E = 271 \angle \beta_1 + 244 \angle \beta_2 + 244 \angle \beta_3$$

Phase Angles of Fields

All we have to do now is find values for the βs. The angle β is actually the relative phase of the signal from each tower when it arrives at point P. (Remember from Chapter 7 that the way we got a desired pattern in the first place was to arrange things so the signals from the various towers would arrive at a point with different phases. The phase angle β of a signal at a point in space depends on the orientation of the towers, the spacing between the towers, and the relative phases of the currents in the towers. It is convenient to think of the angle β as being the sum of two other angles. One is the *space-phasing angle*, which accounts for the orientation and spacing of the towers. The other is the relative phase γ of the current in the tower. Thus

$$\beta = \text{space-phasing angle} + \gamma$$

The space-phasing angle of each tower will have the form

$$\text{Space phasing angle} = S \cos (\theta_{\mathrm{n}} - \theta)$$

where S is the distance in degrees between the tower and the space-reference point of the array, θ_n is the angle between the space-reference axis and a line from the space-reference point to tower n, and θ is the azimuth angle from the space-reference axis to the point P at which we wish to find the field intensity. There will be a space-phasing angle for each tower.

We have assumed in Fig. 8-2 that all of the towers lie on the space-reference axis; therefore, θ_1 and θ_2 are equal to zero. Angle θ_3 is equal to 180° because it is below the reference tower. In Fig. 8-2 we see that tower 1 is at the space-reference point, so S_1 is zero. Inasmuch as tower 1 is our reference tower, γ_1 is also zero. Thus β_1 becomes zero. At tower 2 we find that the space-phasing angle is

$$S_2 \cos(\theta_2 - \theta) = 110 \cos(0 - \theta) = 110 \cos \theta$$

Angle θ in this problem is 40°. Thus

$$S_2 \cos(\theta_2 - \theta) = 110 \cos(0 - 40) = 84°$$
$$\beta_2 = 84° + 5° = 89°$$

Similarly, at tower 3

$$S_3 \cos(\theta_3 - \theta) = 110 \cos(180 - 40) = -840$$
$$\beta_3 = -84° - 5° = -89°$$

Completing and Tabulating the Solution

Now we have numbers for the βs that we can substitute into our equation, which becomes

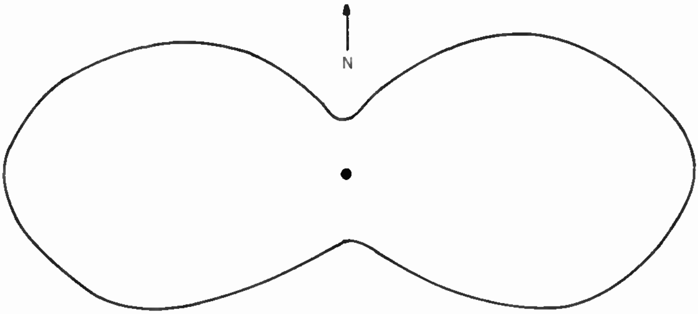
$$E = 271 \angle 0° + 244 \angle +89° + 244 \angle -89° = 280 \text{ mV/m}$$

This tells us that the effective field at point P, which is at an angle of 40°, is 280 mV/m under the conditions described in Fig. 8-2. We could easily use our equation to compute the entire pattern of the array. Letting the azimuth angle ϕ from the space reference be the variable, our now-familiar equation becomes

$$E = 271 \angle 0 + 244 \times \angle 110 (\cos \phi) + 5°$$
$$+ 244 \angle 110 \cos(180 - \phi) - 5°$$

Using a calculator, we can compute the values of the vectors and tabulate them at various angles as in Fig. 8-3.

This procedure can be used to find the effective field at one mile at any azimuth for any number of towers. The technique removes much of the mystery from the operation of a directional-antenna array and is summarized in Fig. 8-4.



ϕ	$244\angle 110(\cos\phi) + 5^\circ$	$244\angle 110\cos(180 \,\iota\, \phi) - 5^\circ$	E (mVm)
0	$244\angle 115$	$244\angle -115$	65
30	$244\angle 100$	$244\angle -100$	186
60	$244\angle 60$	$244\angle -60$	515
90	$244\angle 5$	$244\angle -5$	757
120	$244\angle -50$	$244\angle +50$	585
150	$244\angle -90$	$244\angle +90$	271
180	$244\angle -105$	$244\angle +105$	145

Fig. 8-3. Pattern of array of Fig. 8-3.

Finding the Contribution of Each Tower

When an engineer finds that his antenna pattern is outside of its prescribed tolerances at some azimuth, it is very helpful if he knows just how the magnitude and phase of each of the tower currents contributes to the field intensity at that particular angle. We can find this by drawing a vector diagram for the field intensity at the angle of interest. All that we need for the job is a ruler and a protractor.

Suppose that we wish to study the field intensity of the pattern of Fig. 8-3 at an azimuth of 30°. From the table we find that the equation for the field at 30° is

$$E = 271 \angle 0^\circ + 244 \angle 100^\circ + 244 \angle -100^\circ$$

1. FIND MAGNITUDE OF FIELD VECTOR OF EACH TOWER.
2. FIND CURRENT PHASE ANGLE OF EACH TOWER.
3. ADD TOWER-CURRENT PHASE ANGLE TO ABOVE (2).
4. ADD VECTOR CONTRIBUTION OF EACH TOWER.

Fig. 8-4. Finding pattern of any array.

Inasmuch as we are only interested in the relative contribution of each tower, we can simplify the equation somewhat by dividing through by the field intensity from the reference tower, giving us

$$1E = 1 \angle 0 + 0.9 \angle 100 + 0.9 \angle -100$$

The first step is to draw the vector corresponding to the first term on the right side of the equation. For convenience we will arbitrarily let 1 in. correspond to a relative field intensity of 1. The angle is zero, so we draw a 1 in. line from the origin to the right, as shown in Fig. 8-5A. Next, at the end of this vector, we draw another vector, corresponding to the field intensity from tower 2. The relative field intensity from tower 2 is 0.9 so we make this line 0.9 in. long. The phase angle is 244°, and we consider a positive angle to be a counterclockwise rotation. Therefore, with a protractor, we draw 0.9 in. line from the tip of the first vector at an angle of 100°, as shown in Fig. 8-5B. The final step is to draw the vector for the field intensity from tower 3. This is a 0.9 in. line, drawn from the tip of the second phasor at an angle of −100° from the reference axis. The negative sign means that the angle is measured in a clockwise direction.

The field at 30° is the resultant of the three vectors and is labeled E in Fig. 8-5C. By studying this vector diagram, we can not only determine the field at 30°, we can see just how the fields from the individual towers combine to form the desired field intensity. A vector diagram of this type can be drawn for each azimuth angle of interest. It is a good idea to prepare such a diagram for at least the angles on which the stations licensed monitoring points lie. With these diagrams available the engineer can to some extent determine just how each of the

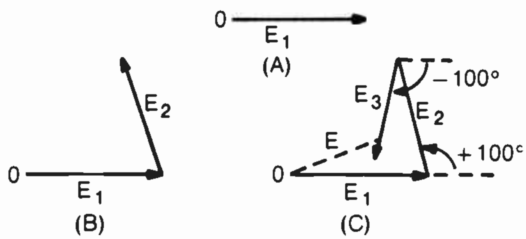


Fig. 8-5. Vector contribution from each of three towers.

parameters of the system influences the field intensity at each of the monitoring points. Then, with this information he can plan the adjustments in advance and avoid a lot of cut-and-try operations.

Contribution of Various Tower Currents and Phases

In a 2-tower array a change in the amplitude or phase in either of the towers has an equal effect at any point, as long as the currents are equal. When an array has three or more towers, the effect of each of the towers is not equal at all bearings. In any case, a vector diagram can show just how a change in any of the tower currents or phases will affect the field intensity at the bearing for which the diagram was drawn.

Figure 8-6 shows a vector diagram for a 3-tower array. This particular diagram was drawn to show the field intensity in a partially filled null. By studying the figure, we can learn a great deal about the factors that affect the field intensity on a particular radial. First of all, we can see how making the current in tower 2 smaller than the other two currents fills in the null. If the vector corresponding to E_2 were as long as the others, the tail of the last vector would land right on the starting point, indicating a field intensity of zero. By making the vector for E_2 a little shorter than the others, we can avoid complete signal cancellation in the null.

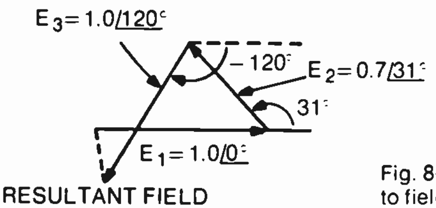


Fig. 8-6. Analysis of contributions to field intensity at any point.

We can also see from the figure how the phases of the currents in towers 2 and 3 affect the resultant field intensity. Lets start with tower 2. If the phase angle was a little smaller, the effect would be to raise the tail of the final vector closer to the origin. Hence reducing the phase angle of tower 2 would reduce the length of the resultant. In other words, it would reduce the field intensity in the null. If we were to increase the magnitude of the current in tower 2, we would raise the tail of the final vector, thus reducing the field intensity in the null.

This type of vector diagram is probably more useful to the broadcast engineer than any other mathematical tool except his calculator. By constructing such diagrams for each radial that has a monitoring point, he can gain a great deal of insight into how the various adjustments of the phasor will affect the field intensity at each of the monitoring points.

Three Nonaligned Towers

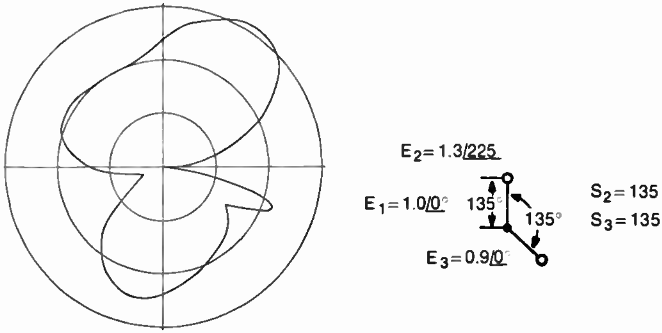
Whenever the towers in a directional-antenna array are in a straight line the field will be symmetrical about the line. This is a great convenience when we are computing the radiation pattern because we only have to compute the pattern on one side of the line of towers; the other half of the pattern will be the same. Unfortunately, it isn't always possible to obtain the pattern required for a particular station location with an in-line array. The requirements for protecting other stations and providing service to the primary service area may well call for a radiation pattern that is not symmetrical at all. This type of pattern is obtained by using towers that are not in line.

Figure 8-7 shows an unsymmetrical pattern produced by a 3-tower array in which the towers are not in line. This arrangement is often called a *dogleg* array.

Once we know the parameters of the array, we can compute the field intensity along any radial or the entire pattern, using the same techniques as for a 3-tower in-line array. Where ϕ is the azimuth angle, we can write the vector equation for the array of Fig. 8-7 by inspection.

$$E = 1\angle 0° + 1.3 \ \underline{/135 \cos \phi + 225°} + 0.9 \ \underline{/135 \cos (135° - \phi)}$$

To keep things simple, we will only consider the *shape* of the pattern from this array. We will therefore let the field intensity from the reference tower equal 1. Solving the preceding equation for each 10° gives the table of field intensities in Fig. 8-7. Note that when the towers in the array are not in line, the pattern is not necessarily symmetrical about the line of towers. We must, therefore, compute the pattern all around the antenna, not merely for one half as we did with the in-line



φ	135(cosφ) + 225°	135 cos(135° ι φ)	E
10	358	−77	2.7
20	352	−57	2.9
30	342	−35	3.1
40	328	−12	3.1
50	312	+12	2.9
60	293	35	2.3
70	271	57	1.6
80	248	77	0.8
90	225	95	0.0
100	202	111	0.6
110	179	122	2.2
120	158	130	1.4
130	138	134	1.6
140	122	134	1.8
150	108	130	1.9
160	98	122	2.1
170	92	111	2.2
180	90	95	2.4
190	92	77	2.5
200	98	57	2.4
210	108	35	2.2
220	122	12	1.8
230	138	−12	1.1
240	158	−35	0.5
250	179	−57	0.8
260	202	−77	1.4
270	225	−95	1.8
280	248	−111	2.1
290	271	−122	2.1
300	292	−130	2.1
310	312	−132	2.0
320	328	−134	2.0
330	342	−130	2.0
340	352	−122	2.0
350	358	−111	2.2
360	360	−95	2.4

Fig. 8-7. Three tower dogleg array with pattern and data.

towers. We can now plot the pattern of this array as shown in Fig. 8-7. Note that there are nulls at azimuth angles of about 90° and 240°. The 240° null is partially filled. The main lobes of the pattern are not symmetrical, hence most of the population of the primary service area would be located north and south of the station site.

The null at 90° is very deep. It would probably be hard to keep the field intensity along this radial within the licensed value. This would be a good radial on which to construct a vector diagram to see just how the magnitude and phase of each of the tower currents contribute to the field intensity. Figure 8-8 shows such a vector diagram, constructed by drawing a vector of the proper length and angle for the field intensity from each tower along the 90° radial. Assuming that one tower of the 3-tower array is located at the space-reference point of the array, both the spacing and angular orientation of the two other towers can be varied by the designer. In addition, the relative magnitude and phase of the two other tower currents can be varied. This permits an extremely large number of patterns to be obtained by design. By the same reasoning, an engineer attempting to merely *adjust* the array can also produce a wide variety of patterns—none of which may be the one that he is trying to obtain. If vector diagrams are drawn to show how the phase and amplitude of each of the tower currents contribute to the field intensity at various points in the pattern, the adjustments will be easier to make. A diagram of this type should be constructed for at least each of the monitoring points specified in the station license. Of course, there will always be some interaction between controls that can't be predicted in this way.

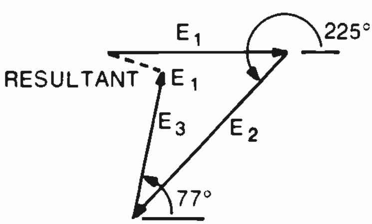
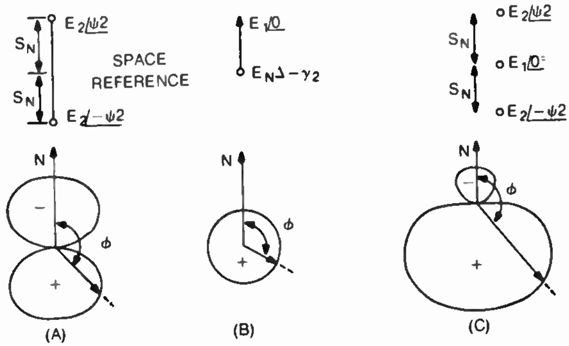


Fig. 8-8. Vector diagram for field intensity along 90° radial of array of Fig. 8-7.



Azinutl	Pattern A	Pattern B	Pattern C
10	−2.00	1	−1
20	−2.00	1	−1
30	−1.96	1	−.96
40	−1.87	1	−.87
50	−1.69	1	−.69
60	−1.41	1	−.41
70	−1.02	1	−.02
80	−0.54	1	0.46
90	0.0	1	1.00
100	0.54	1	1.54
110	1.02	1	2.02
120	1.41	1	2.41
130	1.69	1	2.69
140	1.87	1	2.87
150	1.96	1	2.96
160	1.99	1	2.99
170	2.00	1	3.00
180	2.00	1	3.00
190	2.00	1	3.00
200	1.99	1	2.99
210	1.96	1	2.96
220	1.87	1	2.87
230	1.69	1	2.69
240	1.41	1	2.41
250	1.02	1	2.02
260	0.54	1	1.54
270	0.00	1	1.00
280	−0.54	1	0.46
290	−1.02	1	0.02
300	−1.41	1	−0.41
310	−1.69	1	−0.69
320	−1.87	1	−0.87
330	−1.96	1	−0.96
340	−1.99	1	−0.99
350	−2.00	1	−1.00
360	−2.00	1	−1.00

Fig. 8-9. Synthesis of pattern by addition.

SYNTHESIZING A PARALLELOGRAM ARRAY

In one 4-tower array the towers are located at the corners of a parallelogram. The pattern of this array is obtained by combining the patterns of two 2-tower arrays that are oriented

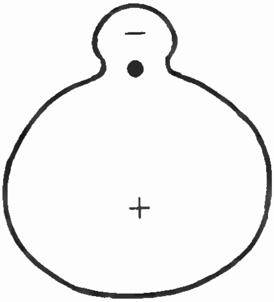


Fig. 8-10. Null filling in 3-tower array.

at different angles. Frequently very complex patterns with many nulls are formed in this way. Each of the two arrays has a tower spacing greater than 180° so each pattern will have more than two nulls.

Parts A and B of Fig. 8-11 show two 2-tower arrays plus their patterns. In Fig. 8-11C the two patterns are added together to form a very complex pattern.

This procedure of adding towers to an array to provide a more intricate pattern can be used to provide just about any type of pattern. A fifth or even a sixth tower can be added to the array of Fig. 8-11 to further change the pattern.

PATTERN SYNTHESIS BY MULTIPLICATION

One common method of synthesizing a desired pattern is to multiply two patterns together. Parts A and B in Fig. 8-12 show the patterns of two 2-tower arrays. If the two patterns

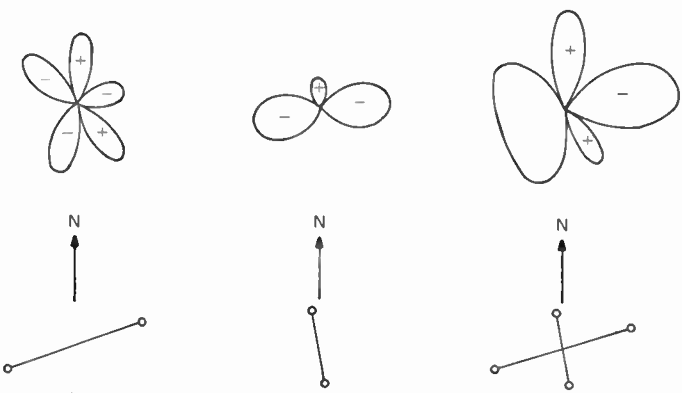


Fig. 8-11. Addition of 2-tower patterns to form a 4-tower parallelogram.

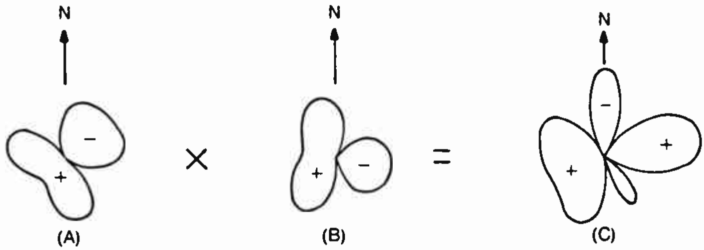


Fig. 8-12. Multiplication of patterns to synthesize a 4-tower array.

are multiplied together, the pattern of Fig. 8-12C results. Since 0 multiplied by any number is still 0, the resultant pattern will have all of the nulls of the component patterns. If both of the component patterns have negative lobes in the same direction, the resultant pattern will have a positive lobe in this direction, since the product of two negative numbers is a positive number. The mathematical development of the resultant field intensities is rather lengthy and thus will not be given here.

The multiplication method has been used frequently in the design of 4-tower parallelogram arrays because of its simplicity. If a station is required to protect the service areas of other stations in four different directions, one 2-tower array can be designed with two nulls and oriented so that it will provide the desired protection in two directions. Then another 2-tower array can be designed to provide protection in the two remaining directions. Finally, the two patterns can be multiplied together to form a 4-tower parallelogram array.

Losses from High-Angle Radiation

A directional-antenna system is designed so that high-angle radiation is minimized and radiation along the surface of the earth is maximized. This increases the field intensity in the primary service area. Depending on the design, the rms field from an array may be either greater or less than the rms field that would be produced by one of the towers radiating the same amount of power. The antenna gain is a measure of the increase or decrease of radiation along the surface of the earth. If the gain is greater than 1, the array will have a greater rms field than would be produced by one of the towers radiating the same amount of power.

The gain of an array is a theoretical consideration and has nothing to do with *losses in the system*. It is defined by the equation

$$G_p = \frac{E_0^{\ 2}}{E_1}$$

where

G_p = power gain in horizontal direction
E_0 = rms horizontal-field intensity when array is radiating full power
E_1 = rms horizontal-field intensity that would be produced with reference tower radiating all of the power

The gain is inherent in an array's design. Some systems have gain over the reference tower, but others have a loss.

DIRECTIONAL-ANTENNA EFFICIENCY

To completely describe the performance of the array, we must introduce another term—the *efficiency* of the array—which takes into consideration the losses in the system. The efficiency η is defined by the equation

$$\eta = \frac{P_r}{P_r + P_l}$$

where P_r = power actually radiated in system
P_l = power lost or dissipated in system, expressed in the same units as P_r

Losses in System

The efficiency of a directional antenna is related to the rms field intensity by the equation

$$\eta = \frac{E_{0L}}{E_0}$$

where E_0 = rms horizontal-field intensity excluding losses
E_{0L} = rms horizontal-field intensity including losses

One design objective is to keep the efficiency as high as practicable by reducing losses in the system. Another is to reduce losses from high-angle radiation.

Chapter 9

Directional-Antenna Impedances

To feed energy to each of the towers of a directional-antenna system with a minimum amount of reflection, the driving-point impedance seen looking into the base of each tower must be matched to the characteristic impedance of the transmission line. Before this can be done, the driving-point impedance of each tower must be known. In Chapter 6 the driving-point impedance of a single tower was seen to depend only on the physical characteristics of the tower and, to some extent, on the ground system. Once the impedance of a single vertical tower is found, it isn't likely to change unless something goes seriously wrong.

When a tower is used as an element of a directional-antenna array, its driving-point impedance usually will not be anything like what it would be if the tower were acting alone. Furthermore, the driving-point impedance changes with the magnitude and phase of its current and the currents of the other towers in the array.

Before going into the details of the impedances found in directional-antenna systems, let's briefly review just what impedance is. Impedance is always the ratio of a voltage to a current. A driving-point impedance between two terminals, such as between the base of a tower and the ground system, is the ratio of the voltage across these terminals to the current flowing in them. Usually the voltage and current will not be in phase, so the impedance will be a vector or complex quantity.

In a directional-antenna system we are interested primarily in the *current* in each of the towers. We adjust the networks in the system so that the tower currents have the proper magnitude and phase, and we let the voltage fall where it may. We are interested in the impedance at the base of each tower because we wish to minimize reflections on the transmission lines and to control the phase shift in the matching networks.

To understand the driving-point impedance at the base of a tower, we must understand the factors that control it. We can consider the base connections of two towers of a directional antenna as being two sets of terminals of a network like that in Fig. 9-1. The equation for the driving-point impedance is

$$Z_1 = Z_{11} + I_2 / I_1 \, Z_{12}$$

where Z_{11} = self-impedance of tower
Z_{12} = mutual impedance
I_1 , I_2 = current in towers 1 & 2

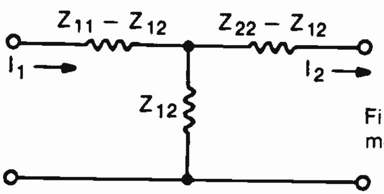


Fig. 9-1. Self-impedance and mutual impedance.

The first term of this equation, Z_{11} , is the self-impedance of the tower. It is what the driving-point impedance would be if there were no other towers close by. The second term consists of the ratios of the base currents in the two towers and the mutual impedance between them. The mutual impedance is often a source of confusion, so we will spend a little time looking into it.

MUTUAL IMPEDANCE

Whenever current flows in a transmitting tower, a voltage is induced in other towers nearby. This complicates the problem of feeding energy to the towers of an array considerably. The ratio of a voltage induced in one antenna—say, tower 1—by the current flowing in another antenna, tower 2, is the *mutual impedance* between the two antennas. Because the system is linear and bilateral, the mutual impedance is the same regardless of which antenna carries the current and which has the induced voltage. The computation of mutual impedance is a mathematical nightmare. Furthermore, mutual-impedance computations are not always accurate. The designer of a system often makes measurements to verify his mutual-impedance computations.

Fortunately, the average broadcast engineer will never be called on to actually compute the mutual impedance between two towers. He should, however, have some idea of what mutual impedance is and how it affects the operation of an antenna system. There are several things we can determine about mutual impedance without actually computing it.

There is no easy way to find the mutual impedance between two antennas; the mechanism by which energy is coupled from one antenna to another is rather complicated. One thing that we can do to help the situation is to redefine the self-impedance of a tower slightly. We can consider the single tower as a device to which we furnish a current and we can think of this current as being the cause of anything that happens in the tower. This is a little different point of view than we normally have, because in circuit theory it is customary to think of voltage as the cause and current as the effect.

If we think of the current in a tower as the cause, we can think of the voltage along the tower as being an *induced voltage* that results from the current. Thus the self-impedance of a tower becomes the ratio of the voltage induced in it to the current that causes the voltage. With the aid of this definition, we can see that if we moved a second tower extremely close to a driven tower, the voltage induced in the second tower would be very nearly the same as the voltage induced in the driven tower. Hence, in this extreme case, the mutual impedance between the towers would be the same as the self-impedance of either tower (assuming that they are identical in construction).

As we move a second tower away from a driven tower, we can safely assume that the induced voltage will be less. We also know that the field will take some time—although an extremely small amount of time—to get from the first tower to the second. Thus there will be some phase shift between the voltage induced in the driven tower and the voltage induced in the second tower, and the mutual impedance will be a vector or complex number. If the induced voltage in the second tower depended only on the radiation field of the first tower, things would be easier. Unfortunately, there is a great deal of coupling between the induction fields of the two towers. The induction field depends heavily on the actual physical configuration of the towers and is not easy to anticipate.

Some idea of the magnitude and phase of the mutual impedance between two towers can be gained from Fig. 9-2. There we have a plot of the resistive and reactive components of mutual impedance between two 90° antennas versus the spacing between them. Some mathematical difficulties were avoided by assuming that the antennas were infinitely thin and that the current distribution was sinusoidal. Thus, the curves cannot be applied directly to practical towers, but they do give a general idea of how mutual impedance behaves.

The first things that we notice about the curves is that both the resistive and reactive components seem to be nearly

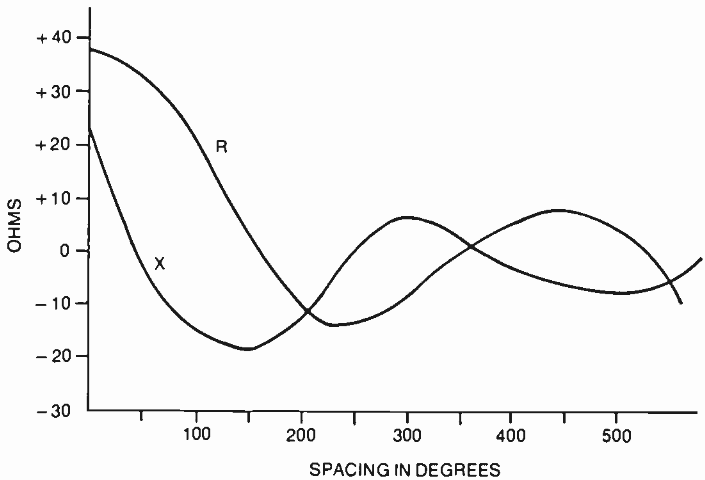


Fig. 9-2. Mutual resistance and reactance between two 90° towers.

periodic, and for some values of spacing they are negative. This isn't very upsetting in the case of the reactive component. Negative reactance is capacitive reactance, and that the curve varies between positive and negative values means simply that the reactive part of the coupling between the two antennas may be either inductive or capacitive, depending on the spacing between them.

That the resistive component of the mutual impedance becomes negative for some values of antenna spacing tends to be disconcerting at first. Actually, a negative resistive component merely means that at some spacings the phase of the induced voltage has changed 180°. Inasmuch as the phase of the induced voltage depends on the spacing between the antennas, it is understandable that it should be negative at some spacings. In fact, the magnitude of the mutual impedance would be nearly periodic except that the resistive and reactive components do not reverse polarity at the same spacings.

Another factor of interest in the curves of Fig. 9-2 is that the magnitude of both the resistive and reactive components of the mutual impedance tends to become larger as the two antennas are placed closer to each other. This is logical, because when the two antennas are closer together, more energy is coupled from one to the other.

When two towers are of equal height we can simplify the situation somewhat, because the magnitude of the mutual impedance at any given spacing is directly proportional to the radiation resistance of each of the towers. The constant of proportionality is a function of the spacing between the antennas, as is the angle of the mutual impedance. Figure 9-3 shows a plot of the ratio of the magnitude of the mutual impedance to the radiation resistance R_r of the towers. This plot is based on infinitely thin antennas, but is usually within about 5% of the actual value. Also shown in the figure is a plot of the angle of the mutual impedance as a function of tower spacing. The angles are not as accurate as the ratios, because the reactive component of the mutual impedance is dependent on the induction fields of the antennas. The curves are nevertheless close enough for the average broadcaster's use.

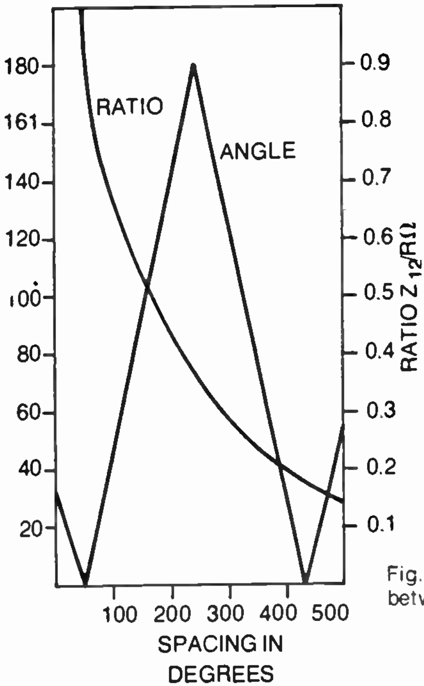


Fig. 9-3. Mutual impedance between towers of equal height.

In practice, mutual impedance is often measured in addition to being calculated. Once it has been measured, it rarely will change very much unless radical changes have taken place in the system. Usually, anything that changes the mutual impedance between two towers will also change the current distributions along the towers. This type of change might be a change in guy wires or some fault in tower-lighting lines. Changes of this sort can usually be detected from base-impedance measurements, particularly when they are made at more than one frequency. Whenever any change is noted in base impedances, the ground system is also suspect.

DRIVING-POINT IMPEDANCE

The impedance that we are most interested in when dealing with a directional-antenna system is the driving-point impedance at the base of each tower. This is the impedance that we must match to the characteristic impedance of the transmission line. The equation

$$Z_1 = Z_{11} + I_2 / I_1 \; Z_{12}$$

for driving-point impedance, shows the various things that will influence the magnitude and angle of this impedance. One factor that will have a constant influence is the self-impedance of the tower. This depends only on the physical characteristics of the tower itself, and not on currents in the system. The second term of the equation causes all of the problems. There are several things that we can deduce about the influence of this term without bothering to put numbers into the equation.

In a wide-spaced array the mutual impedance is small, and there will not be as much interaction between towers as in a close-spaced array, where the mutual impedance is large. Also, if the second term of the equation is large, it will have a pronounced effect on the value of the driving-point impedance.

The examples and equations given so far have been concerned only with the driving-point impedance of a 2-tower array. In many directional-antenna systems there are more than two towers. The driving-point impedance seen looking into the base of a tower in a multitower array is given by

$$Z_1 = Z_{11} + \frac{I_2}{I_1} Z_{12} + \frac{I_3}{I_1} Z_{13} + \frac{I_4}{I_1} Z_{14} + \ldots \frac{I_n}{I_1} Z_{1n} + \ldots$$

This is the same type of equation presented earlier, expanded for systems of any number of towers n. All of the mutual impedances between the tower of interest and each of the other towers, as well as the magnitude and phase of each of the other tower currents, have an influence on the driving-point impedance of each tower in the system. The amount of influence depends on the relative size and phase of each term in the equation.

This is obviously a rather complex state of affairs. There are now so many different parameters that influence each driving-point impedance that it can take on nearly any value. The driving-point impedance of each tower can be calculated, although the calculations tend to be tedious. We will consider these calculations later in this chapter. Here we are concerned with the general values that the driving-point impedance might have, as well as the physical significance that it might have.

Before transforming the driving-point impedance of a tower to the characteristic impedance of a transmission line, we first tune out the reactive component as shown in Fig. 9-4. If the reactive component of the driving-point impedance is inductive, the tuning reactance in Fig. 9-4 will be capacitive; similarly, if the reactive part of the driving-point impedance is capacitive, the tuning reactance will be inductive. Thus, as far as the impedance-matching network is concerned, we are only interested in the resistive part of the driving-point impedance. The reactive part will be tuned out. The resistive part is what is transformed to the characteristic impedance of the transmission line.

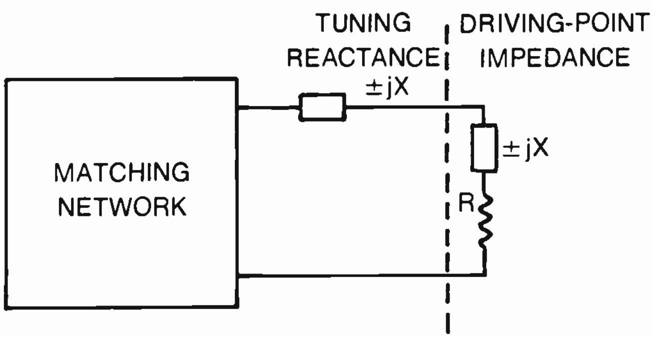


Fig. 9-4. Reactive part of driving-point impedance is tuned out.

This doesn't mean that the reactive component of a driving-point impedance isn't important; it is. If the reactive component is very large compared to the resistive component, as it is in the case of very short towers, the impedance will be hard to match, and the arrangement will tend to have a very narrow bandwidth.

The resistive part of the driving-point impedance of a tower might have almost any value—positive, negative, or even zero. The positive resistance is easy to understand. It is simply the ratio of voltage to current at the output of the impedance-matching network. A positive resistance means that energy is flowing into the tower and is not returning.

Negative Resistance

It is easy to see how the resistive part of the driving-point impedance may be negative. In fact, in arrays containing four or more towers, the driving-point impedance of one or more of the towers is negative much more frequently than the average broadcast engineer would wish. The biggest trouble with negative driving-point resistances is the confusion that they usually produce in the mind of the engineer. Most of the impedances that he works with are positive, and there seems to be something mysterious about a negative impedance or resistance.

Figure 9-5A shows a "black box" with the voltage and current at its terminals. Using Ohms's law, we decide that whatever is in the box is equivalent to a 20-ohm resistor. Because of the polarity of the voltage and the direction of current, we know that energy is flowing into the box.

In Fig. 9-5B we have a similar box with the same magnitudes of voltage and current, but the direction of current has been reversed. Again applying Ohm's law, we decide that whatever is in the box is equivalent to a 20-ohm resistor; but because of the polarity and current direction, we conclude that the resistance is negative. This merely means that the energy is flowing out of the box instead of into it.

$$R = \frac{+10V}{+0.5A} = +20 \text{ OHMS} \qquad R = \frac{+10V}{-0.5A} = -20 \text{ OHMS}$$

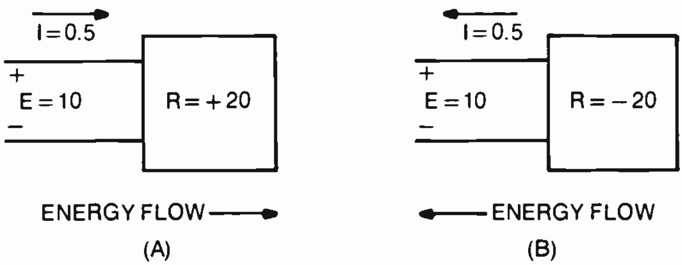


Fig. 9-5. Positive and negative driving-point resistances.

The concept of negative resistance will be a little clearer if we look at a block diagram of a somewhat simplified directional-antenna system (Fig. 9-6). Here all of the energy comes originally from the transmitter. It does not, however, flow directly through the system to the far field. As we saw earlier, some of the energy from each of the towers is coupled to each of the other towers. In two of the towers in the figure, more energy enters the tower through the network than is picked up from the other towers. Thus the impedance at their bases is positive and energy is flowing into them. With the third tower, however, the situation is just the opposite. Here the current is flowing out of the tower rather than into it so the sign of the resistance is negative.

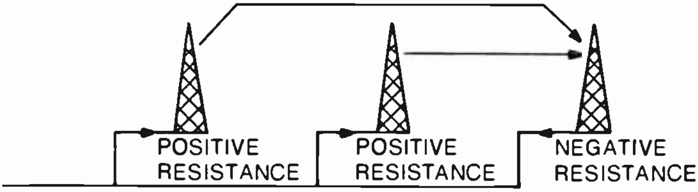


Fig. 9-6. Energy flow into a negative-resistance tower.

One source of confusion regarding negative resistance is that only a driving-point resistance can be negative. Figure 9-7 shows two boxes connected together. From the voltage and current between them, we can conclude that the impedance seen at the input of the second box is a resistance of 20 ohms. Now suppose that the box does in fact contain a 20-ohm resistor, and nothing more, as in Fig. 9-7B. Then, no matter what happens to be in the first box, the impedance seen at the input of the second box will always be a 20-ohm resistance. The reason is that, in this special case, the driving-point impedance seen looking into the second box is also the self-impedance of the 20-ohm resistor. If the second box contained a source, such as a battery, its driving-point impedance would depend not only on what was in the box but on what was connected across its terminals.

Figure 9-7C shows the same two boxes, but this time with batteries and resistors. The impedance measured at the input of the second box is still 20 ohms. If we were to specify this same impedance looking back into the first box, it would still be 20 ohms, but the sign would be negative because energy is flowing out of the first box. The magnitude of the resistance depends on the voltages of the batteries in both boxes, as well as on the values of the resistors they contain. Thus a driving-point impedance is merely the ratio of a voltage to a current at the input of a network or circuit or at the base of a tower. Its magnitude depends on what is connected to it. Its sign depends on the direction in which energy is flowing. There is no point in asking *what* we are measuring the resistance of when we specify 20 ohms in Fig. 9-7C; we are merely specifying the ratio of a voltage to a current. Its value depends on many factors. The same is true when we specify the base impedance of a tower in a directional-antenna system; we are merely specifying a ratio of the voltage to the current at the base. The value of this ratio depends on many different things, including not only the self-impedance of the tower but also the mutual impedances to other towers, as well as all of the currents in the system. Again, if the base impedance is a negative number, this merely means that energy is flowing out of a tower instead of into it.

Zero Resistance

From our equation we can see that not only may the resistive part of the driving-point impedance of a tower be positive or negative, it may very well be zero. Zero resistance is probably even more confusing than negative resistance. What does it mean? If the resistance between two terminals in a circuit is zero, then we have a short circuit, and in general it isn't advisable to apply any voltage at all to a short circuit. Perhaps we can get a better idea of what a zero driving-point resistance means by writing an equation for it,

$$0 = Z_{11} + \frac{I_2}{I_1} Z_{12}$$

Inspection of the above equation shows that what it really means is that no voltage at all is required to make the current I_1 flow. Although this situation is impossible in passive circuits employing only real resistors, it is not at all uncommon in antennas. It means that the element whose driving-point impedance we are considering is *parasitic*; the current is caused by the voltage induced in the element by the currents in the other elements of the array. All we have to do is to short its terminals, and the proper amount of current will

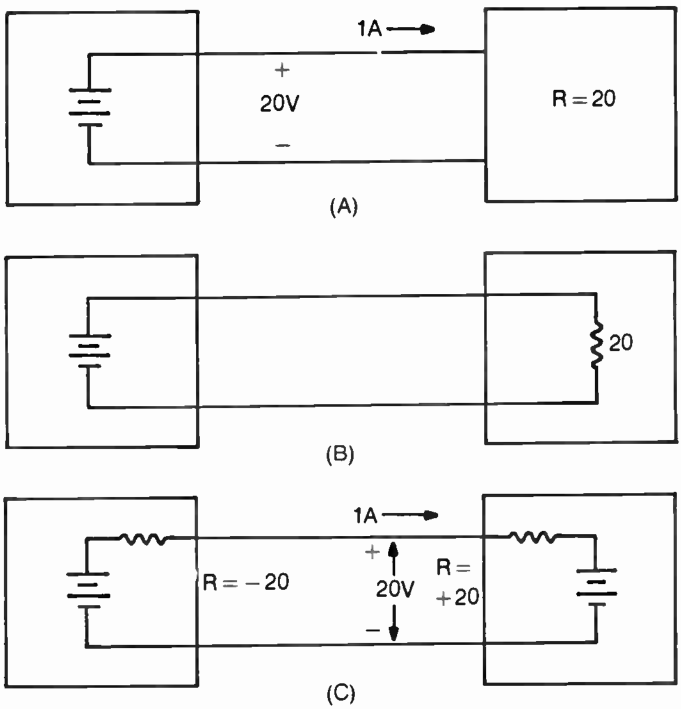


Fig. 9-7. Driving point resistances discussed in text.

flow to make the element behave as it should. This is true of the directors and reflectors in a Yagi antenna.

At first glance, it might appear that a zero driving-point impedance would be desirable. It would save on transmission lines. All that we would need at the base of the tower would be a network to tune out the reactance. In general, however, the practice isn't followed in broadcast-antenna systems. The reason is that even the slightest change in the parameters of the system will make the driving-point impedance go either positive or negative, changing the parameters of the feeder system. Usually, whenever it is possible, the use of towers that have a very low or zero driving-point resistance is avoided. When such a tower is unavoidable in the design of an array, it is quite common to add some series resistance at the feed point. This will introduce some loss into the system; but usually there isn't much power fed to such a tower, and the improvement in the system is well worth the additional loss.

COMPUTING DRIVING-POINT IMPEDANCE

The actual computation of a driving-point impedance has traditionally been considered a very tedious operation. It is necessary to convert several impedances from polar to rectangular form and back again. Fortunately, with the advent of a pocket calculator that will handle vector addition, this in no longer necessary. This capability alone will make the calculator a very worthwhile investment for the broadcaster who is involved in the computation of driving-point impedances.

Let us consider as an example a 2-element array that has the following parameters:

$$Z_{11} = 25 + j40 \qquad Z_{22} = 25 + j40 \qquad Z_{12} = 20 - j15 \qquad \frac{I_2}{I_1} = 0.9 \angle 100^\circ$$

The first step is to express everything in polar form as follows:

$$Z_{11} = 47 \angle 58^\circ \qquad Z_{22} = 47 \angle 58^\circ \qquad Z_{12} = 25 \angle -37^\circ \qquad \frac{I_2}{I_1} = 0.9 \angle 100^\circ$$

Now we can proceed to compute the driving-point impedance of tower 1 by the equation

$$Z_1 = Z_{11} \quad \frac{I_2}{I_1} Z_{12}$$

Substituting numbers into this equation gives us

$$Z_1 = 47 \angle 58^\circ + 0.9 \angle 100^\circ + 25 \angle -37^\circ$$
$$Z_1 = 47 \angle 58^\circ + 22.5 \angle 63^\circ$$
$$Z_1 = 65 \angle 60^\circ$$

Now we convert back to rectangular form and find the driving-point impedance of tower 1 to be

$$Z_1 = 35.12 + j60 \text{ ohms}$$

We can find the driving-point impedance of tower 2 similarly. The only difference in the two equations is in the second term;

the ratio of the currents is now 1.11 instead of 0.9. Hence we have

$$Z_2 = Z_{22} + \frac{I_1}{I_2} Z_{12}$$

$$Z_2 = 47 \angle 58^\circ + 1.1 \angle -100^\circ + 25 \angle -37^\circ$$
$$Z_2 = 47 \angle 58^\circ + 28 \angle -137^\circ$$
$$Z_2 = 22 \angle 78^\circ$$

The driving-point impedance of tower 2 is thus

$$Z_2 = 4.5 + j22$$

The computation of the driving-point impedances is clearly just a matter of pressing a few keys on a calculator. When more than two towers are involved, there are a few more steps; but the process isn't at all difficult. In fact, if the broadcast engineer would go through the steps, whether it is necessary or not, he would gain a great deal of insight into just how his antenna really works and what interaction between ratio and phase controls to expect.

Chapter 10

Impedance Transformation and Phase Shifting

The average broadcast engineer finds the networks used in the feeder system of a directional antenna confusing at best. Most textbook treatments of networks are highly mathematical—more so than is needed for maintaining and operating an antenna system. The few nonmathematical treatments usually make heavy use of graphs that are hard to understand and even harder to interpret. The subject is actually more unfamiliar than difficult. The small amount of effort required to understand how the networks in a feeder system operate is well worth while. The better an engineer understands how the networks in his antenna system function, the less apt he is to get in trouble while adjusting them.

In a directional-antenna system, networks are used for the following purposes:

1. Impedance matching between various parts of the system
2. Phase shifting
3. Power division

Unfortunately, the three functions listed above are not as distinct as one might wish. Impedance-matching networks and power dividers introduce phase shift. Probably the greatest cause of difficulty in adjusting networks is the interaction that takes place between their controls. For example, when one of the controls on a power divider is moved it might cause more phase shift than power division. Once one is familiar with the principles of network operation, this interaction can be anticipated and will not be as formidable.

The subject of power division is treated separately, in the next chapter.

IMPEDANCE-TRANSFORMING NETWORKS

At audio frequencies, impedance matching is usually accomplished by a transformer. This impedance transformation is easy to understand: Impedance is a ratio of voltage to current, and a transformer changes this ratio.

At broadcast frequencies, impedance matching is accomplished by reactive networks. Just how a reactive network makes one value of impedance look like another value is not always clear. We will approach the subject by means of equivalent circuits. Once the concept of equivalent reactive circuits is well understood, the action of various impedance-transforming networks will fall logically into place.

Equivalent Circuits

Figure 10-1A shows a "black box" that contains only passive elements such as resistances and reactances. By

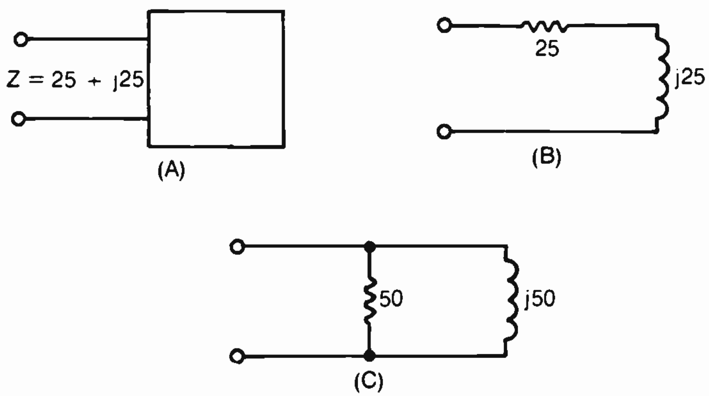


Fig. 10-1. Equivalent series and parallel circuits.

means of an impedance bridge we find that the impedance seen looking into its terminals is

$$Z = 25 + j25 \text{ ohms}$$

Remember, this impedance is a property of whatever happens to be in our box, and it tells us what the magnitude and phase angle of the current will be if we apply a given voltage to the terminals. Likewise, it tells us the magnitude and phase of the voltage that would appear across the terminals if we forced a given current through them. In other words, the impedance at the terminals of the network is a measure of the ratio of the voltage across the terminals to the current flowing in them.

Series and Parallel Equivalence

One thing the box might contain is the simple series circuit shown in Fig. 10-1B, which consists of a 25-ohm resistance in series with a 25-ohm inductive reactance. The important thing to remember is that this isn't the only circuit that our box might contain. For example, it might contain the circuit shown in Fig. 10-1C, which consists of a 50-ohm resistance in parallel with a 50-ohm reactance. On the surface these two circuits don't look at all alike, but a few simple calculations will show that the impedance seen looking into their terminals is exactly the same.

The impedance seen looking into the terminals of the circuit of Fig. 10-1C is found by taking the product over the sum, just as we would do with parallel resistors.

$$Z = \frac{50\,(j50)}{50 + j50}$$

We can rationalize the denominator by multiplying both the numerator and denominator by the conjugate of the denominator.

$$Z = \frac{50\,(j50)}{50 + j50}\;\frac{(50 - j50)}{(50 - j50)} =$$

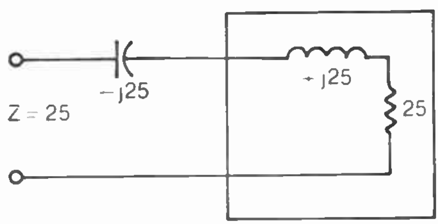
$$\frac{50\,(50^2)}{50^2 + 50^2} + j\,\frac{50\,(50^2)}{50^2 + 50^2} = 25 + j25 \text{ ohms}$$

Thus the impedance seen looking into the circuit, 25 + $j25$ ohms, is exactly the same as the impedance seen looking into the series circuit of Fig. 10-1B.

This equivalence of series and parallel circuits is the basis of all reactive impedance-transforming networks and must be clearly understood. Stated differently, a resistance and reactance connected in series look exactly like a different value of resistance in parallel with a different value of reactance. We can't tell the difference from any measurements that we make at the terminals of the network. This equivalence only holds true if we keep the frequency constant: but this restriction won't cause any trouble, because broadcast stations operate at a constant frequency.

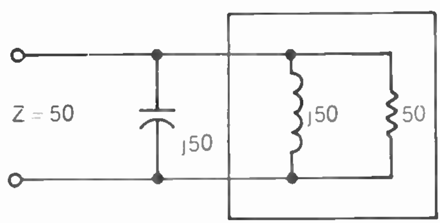
It doesn't make any difference at all to us which circuit is actually in the box of Fig. 10-1A. If we add the proper reactance at the input terminals to tune out the reactance in the box, we can make it look like *either* a 25-ohm resistance or a 50-ohm resistance, depending on how much reactance we connect at the input and how we connect it.

Figure 10-2A shows our black box with a 25-ohm capacitive reactance connected in series with it. The inductive and capacitive reactances cancel, and the impedance seen looking into the circuit is a pure resistance of 25 ohms.



(A)

Fig. 10-2. Two values of pure resistance obtained from the same "black box."



(B)

In Fig. 10-2B we have the same black box, but this time with a 50-ohm capacitive reactance in parallel with the input terminals. Again the inductive and capacitive reactances cancel, and the impedance seen looking into the box is a pure resistance of 50 ohms.

This concept is apt to be confusing at first. Let's go over it again. In Fig. 10-3 we have two terminals between which we measure impedance and find it to be $25 + j25$ ohms. Let's say that we have no idea what is actually in the box. It might be either the series or the parallel circuit that we have considered. Or it might be neither of these; it might be the impedance measured across the base insulator of a tower. In this case, the resistance and capacitance are actually distributed along the tower, and one equivalent circuit is as good as the other. We might tend to think of the circuit in A of Fig. 10-3 as being more "real" than the one in B. But if we had started out with an admittance bridge, we would have found the admittance across the terminals of our box (Fig. 10-3B) to be

$$Y = 0.02 - j0.02 \text{ mho}$$

This is the equation for a 20 mmho conductance in parallel with a 20 mmho inductive susceptance. If we were to state this in terms of resistance and reactance, it would represent a 50-ohm resistance in parallel with a 50-ohm reactance.

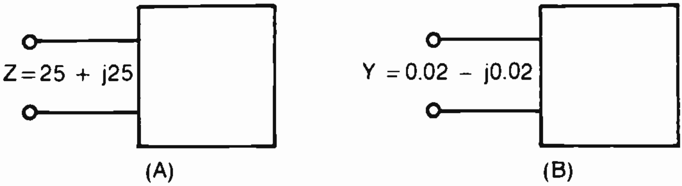


Fig. 10-3. Equivalent impedances and admittances.

Again we see that it makes absolutely no difference what the actual circuit in the box might be. We are only concerned with what it looks like when viewed through its input terminals, and it can be represented equally well by either its series or parallel equivalent.

We can now generalize our equivalent circuits and derive simple equations that will enable us to find either circuit from the other. Before we do this, we should note two things. First of all, there is nothing sacred about the choice of inductive reactance in the example we have been considering. The techniques would work just as well if the reactance inside the box were capacitive; we would merely use an inductive reactance outside the box. In general, impedance-transforming networks use both inductive and capacitive reactance. Second, either a capacitive or an inductive reactance will vary with frequency, so the series and parallel circuits will only be exactly equivalent at one frequency. The two circuits will be approximately equivalent over the bandwidth of a regular broadcast signal, so we can use them in broadcast networks.

Now let's look at our equivalent series and parallel circuits and see if we can use them to understand impedance-transforming networks in general.

Conversion Between Series and Parallel Circuits

Figure 10-4 shows a series network and its parallel equivalent. The parameters of the series network are designated by an *s*, and those of the parallel network are designated by a *p*. We can define the *Q* of the networks in the conventional way. The *Q* of the series circuit is

$$Q = \frac{X_s}{R_s}$$

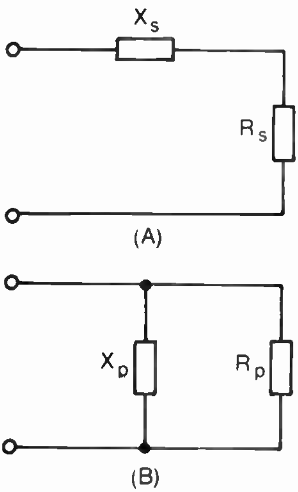
and that of the parallel circuit is

$$Q = \frac{R_p}{X_p}$$

Since the two networks are equivalent, they have the same value of *Q*.

$$Q = \frac{X_s}{R_s} = \frac{R_p}{X_p}$$

If the networks are to be equivalent, the impedance seen looking into their terminals must be equivalent. The



$$X_s = QR_s \qquad R_s = \frac{R_p}{Q^2 + 1}$$

$$X_p = \frac{R_p}{Q} \qquad R_p = (Q^2 + 1)R_s$$

$$Q = \frac{X_s}{R_s} = \frac{R_p}{X_p} = \sqrt{\frac{R_p}{R_s} - 1}$$

$$\frac{R_p}{R_s} = Q^2 + 1$$

Fig. 10.4 Driving equivalent series and parallel circuits.

magnitude of the impedance seen looking into the series circuit is

$$Z_s = R_s + jX_s$$

The magnitude of the impedance seen looking into the parallel circuit is

$$Z_p = \frac{R_p\ (jX_p)}{R_p + jX_p}$$

Now, inasmuch as our two circuits are equivalent by definition, the impedances seen looking into them are also equal. Hence

$$Z_s = Z_p{}^2 = R_s + jX_s = \frac{R_p\ (jX_p)}{R_p + jX_p}$$

We can substitute Q into this equation and simplify it, giving

$$\frac{R_p}{R_s} = Q^2 + 1$$

This basic equation is all that we need to find one circuit when the other is given.

In Fig. 10-4 all of the equations relating the parameters of the equivalent series and parallel circuits are summarized. We can use these equations to design a network that will make any value of resistance look like any other value of resistance as far as we can tell by any measurements. In general, we use one of two procedures.

1. If we wish to make the resistance at the input of a network look larger than the resistance we have, we add a reactance in series, then tune out the equivalent reactance by adding a parallel reactance of the opposite type. Thus, if we added an inductive reactance in series with the existing resistance, we would use a capacitive reactance to tune it out.
2. If we wish to make a resistance look smaller than it is, we add a reactance in parallel with it, then a series reactance of the opposite type to tune it out.

L-NETWORK

The circuits we have been using are actually *L*-networks, circuits whose configuration resembles the letter *L*. By a proper choice of components, such circuits can, in theory, effect any impedance transformation desired. We will see a little later that there are practical limits to the impedance-transformation ratio that we can use in broadcast work.

We will start our consideration of uses for *L*-networks by transforming one value of resistance into another. Later we will consider cases where the impedance is not a pure resistance.

Impedance Transformation by the L-Network

Suppose, for example, that the impedance at the base of a tower is a pure resistance of 10 ohms and that we wish to feed the tower with a 50-ohm transmission line. For maximum power transfer and minimum reflections, we want a network that will transform the impedance of the tower to a pure resistance of 50 ohms. We will start by adding enough reactance X_s in series with the 10-ohm resistance (Fig. 10-5A) to make it look like a 50-ohm resistance in parallel with a reactance X_p (Fig. 10-5B).

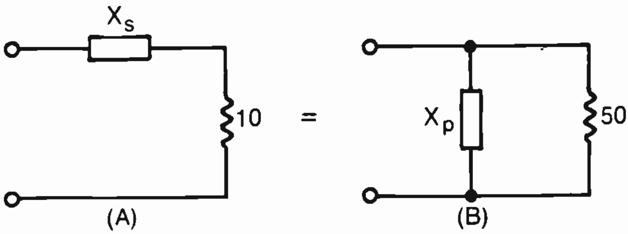


Fig. 10-5. Making a load resistance look larger.

The next step is to find the Q of the network, which is given by

$$Q = \sqrt{\frac{R_p - 1}{R_S}} = \sqrt{\frac{50 - 1}{10}} - \sqrt{4} = 2$$

It is important to note that the Q of the network is determined entirely by the impedance-transformation ratio. Once we decide on the desired ratio, the Q of our network is established and there is nothing we can do about it. This has some rather important implications and puts a limit on the impedance-transformation ratio that we can get in a practical network.

The equivalent series reactance required in Fig. 10-5A is given by

$$X_s = R_s Q = 10 \times 2 = 20 \text{ ohms}$$

The effective parallel reactance of the equivalent circuit of Fig. 10-5B is

$$X_p = \frac{R_p}{Q} = \frac{50}{2} = 25 \text{ ohms}$$

Now we can represent the network we are building by the circuit on the left in Fig. 10-6A. The circuit on the right is equivalent to our network. We can take advantage of this equivalence by adding a 25-ohm capacitive reactance across the terminals of the network (or across the 25-ohm inductive reactance of the equivalent circuit), as shown in Fig. 10-6B, canceling out the reactances but leaving us with a resistance of 50 ohms across the network terminals. We have thus transformed a 10-ohm resistance into a 50-ohm resistance with an *L*-network.

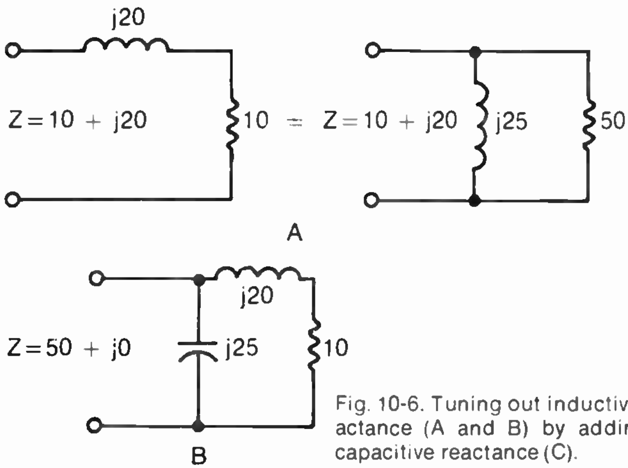


Fig. 10-6. Tuning out inductive reactance (A and B) by adding a capacitive reactance (C).

Now, for another example, suppose that we have a resistance of 100 ohms and wish to transform it into a resistance of 50 ohms (see Fig. 10-7). Inasmuch as we know the required impedance-transformation ratio, we can determine the required Q of the network.

$$Q = \sqrt{\frac{R_p}{R_s} - 1} = \sqrt{\frac{100}{50} - 1} = 1$$

Now, since we are trying to lower the impedance, we will connect a reactance in parallel (Fig. 10-7B). The value of the parallel reactance is given by

$$X_p = \frac{R_p}{Q} = \frac{100}{1} = 100 \text{ ohms}$$

We could use either type of reactance, but we will use capacitive reactance this time. Now we have to find the value of capacitive reactance in the equivalent series circuit of Fig. 10-7D. It is given by

$$X_s = R_s Q = 50 \times 1 = 50 \text{ ohms}$$

This is the value of capacitive reactance that we must tune out of our circuit to have an impedance that is a pure resistance. Hence we will use an inductive reactance of 50 ohms for tuning out the reactance. Our final circuit is shown in Fig. 10-7E. In this case, we transformed a 100-ohm resistance into 50 ohms.

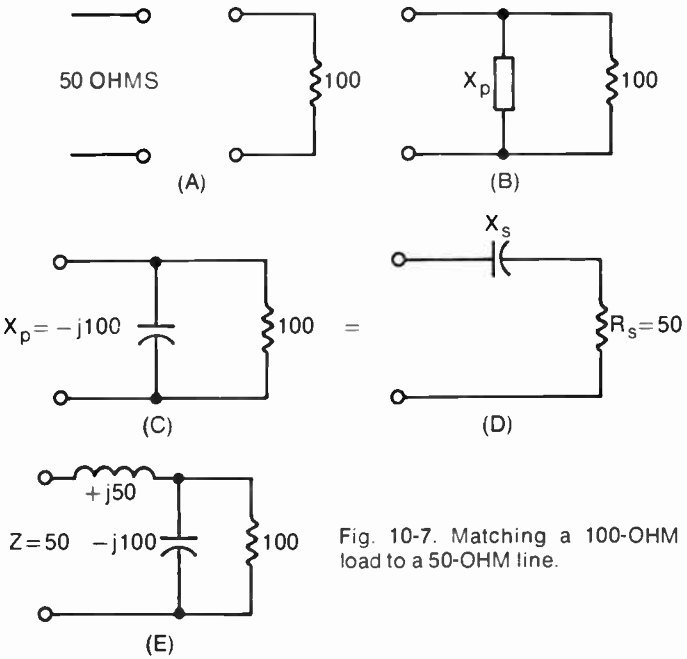


Fig. 10-7. Matching a 100-OHM load to a 50-OHM line.

Configurations of L-networks

Inasmuch as we can use either inductive or capacitive reactance in our *L*-networks, there are two configurations that can be used to transform impedance to a higher value (Fig. 10-8A and B) and two that will transform to a lower value (C and D).

The choice of network configuration isn't arbitrary; we often take advantage of the characteristics of a particular configuration. For example, the circuits of Fig. 10-8A and C act as low-pass filters which tend to reduce harmonic radiation. The circuits of Fig. 10-8B and D act as high-pass filters. Although they are as effective as the others as far as impedance transformation is concerned, they provide no attenuation of harmonics.

There is still another consideration in the selection of an *L*-network configuration. The circuits of Fig. 10-8A and C retard the phase of the current, causing the current at the output to lag the current at the input. The circuits of B and D advance the phase of the current. Sometimes in a directional-antenna system the choice of a configuration is

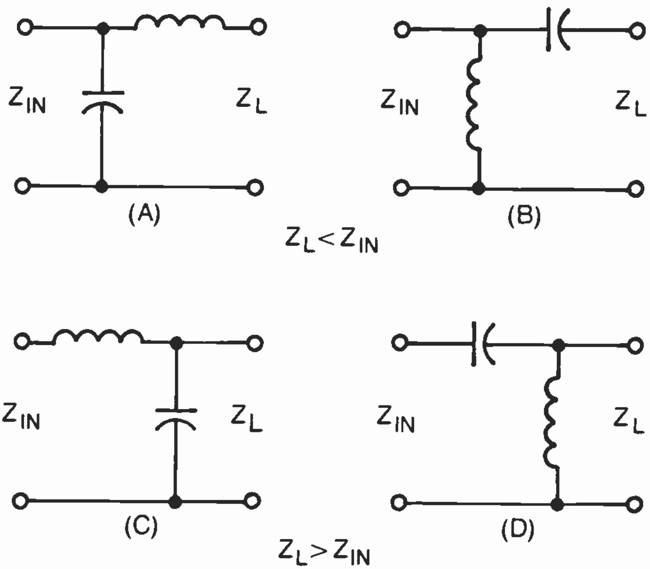


Fig. 10-8. Possible L-network configurations.

decided by a need to introduce phase shift in a line going to a particular tower.

Phase Shift in L-Network

The *L*-network is composed of reactive elements, which store energy; and energy storage takes time. This, in turn, means that there will be some phase shift in the network. Of course, inasmuch as the input and load impedances are pure resistances, the input current will be in phase with the input voltage, and the output current will be in phase with the output voltage. The phase angle exists between the input current and the load current, or between the input voltage and the load voltage.

The amount of phase lead or lag that is introduced by an *L*-network is given by

$$\gamma = \cos^{-1}\frac{1}{\sqrt{r}}$$

where γ = phase shift in degrees
r = ratio of R_L to R_{in}, or R_{in} to R_L, whichever gives $r>1$

This shows that the phase shift, like the *Q* of the network, is completely determined by the impedance-transformation ratio *r*. Once we have selected an impedance-transformation ratio, we have specified the phase shift that the network will introduce, and there is little that we can do about it. We can decide whether we want the phase angle to be leading or lagging and select the network configuration accordingly, but we can't do anything about the magnitude of the phase shift. This is logical when we stop to think that in matching impedances we have to handle the resistance and reactance with the two variable circuit elements. We have nothing left to vary to control the phase shift.

In nondirectional-antenna feeder systems phase shift isn't a problem. We have only one antenna and one transmission line, and we don't care how much phase shift is introduced by our networks. With a directional-antenna array, however, we care very much about the phase of the current feeding each tower. There are many places in a directional antenna where we can't use an *L*-network, because we can't tolerate the phase shift that it would introduce. This doesn't mean that we can't use *L*-networks at all in a directional-antenna system. These networks are, in fact, widely used where the phase shifts in the entire directional-antenna system are such that the inevitable phase shift in an *L*-network can be tolerated.

DESIGN OF L-NETWORK

The values of the reactive components of the generalized *L*-network of Fig. 10-9 can be determined from the equations given. By letting R_1 be the larger of the line or load

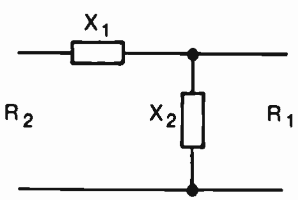


Fig. 10-9. Equations for design of L-network.

$$r = \frac{R_1}{R_2} \qquad a = \frac{r}{\sqrt{r-1}} \qquad b = \sqrt{r-1}$$

$$X_1 = \pm j\,\frac{R_1}{a} \qquad X_2 = \pm j\,\frac{R_1}{b} \qquad \cos\psi = \frac{1}{\sqrt{r}}$$

impedances and R_2 the smaller, the circuit of Fig. 10-9 can represent any of the circuits of Fig. 10-8. Remember that the shunt reactance is always directly across the higher impedance.

It is convenient to express the values of the parameters of the *L*-network in terms of the impedance-transformation ratio *r*, which is given by

$$r = \frac{R_1}{R_2}$$

It is also useful to define two other network parameters, *a* and *b*, in terms of the transformation ratio *r*. These two parameters are given by

$$a = \frac{r}{\sqrt{r-1}}, b = \sqrt{r-1}$$

Of course, these values can be computed easily with an electronic calculator. In case a calculator with square-root capability is not available, the curves of Fig. 10-10 give the

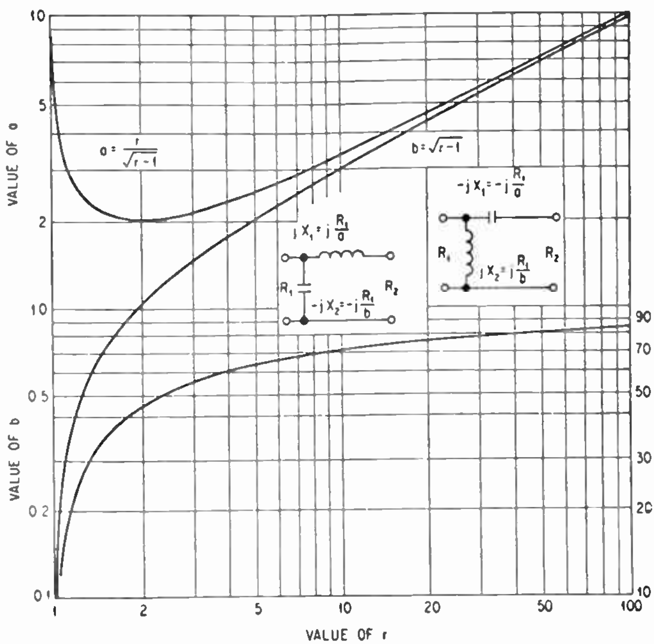


Fig. 10-10. Data for L-network design.

values of a and b in terms of the impedance-transformation ratio r.

Knowing a and b, we can easily find the values of X_1 and X_2 from the equations

$$X_1 = \pm j \frac{R_1}{a} \quad X_2 = \mp j \frac{R_1}{b}$$

The $\pm$ and $\mp$ signs merely mean that if X_1 is an inductive reactance, X_2 must be a capacitive reactance, and vice versa. For example, if we wish to match a 10-ohm load to a 50-ohm transmission line, we find from Fig. 10-10

$$a = 2.5, b = 2$$

Therefore

$$X_1 = j \frac{R_1}{a} = j \frac{50}{2.5} = j20 \text{ ohms}$$

and

$$X_2 = -j \frac{R_1}{b} = -j \frac{50}{2} = -j25 \text{ ohms}$$

This is the same example that is given in Fig. 10-4.

LIMITATIONS OF L-NETWORK

In practice, there is a definite limit to the impedance-transformation ratio that we can obtain with a single *L*-network without running into problems. One limitation of the *L*-network is that it will introduce phase shift as mentioned earlier, and there is nothing that we can do about it. If the system won't tolerate the phase shift, we must use another type of network.

Another limitation of the *L*-network is its *bandwidth*. Our derivation of the *L*-network was based on the use of two reactances, one capacitive and one inductive. The reactance of any real inductance or capacitance changes with frequency, so the inductances and capacitances in a real network exhibit their design values of reactance *at only one frequency*. At frequencies above and below the frequency for which the network was designed, its performance deviates somewhat from the ideal. For broadcast applications the signal transmission through the network must be reasonably constant over the bandwidth of the transmitted signal, otherwise sideband power will be lost.

In Fig. 10-11 the response of an *L*-network having a low-pass-filter configuration is plotted as a function of normalized frequency. The response at the design frequency is given at 1.0 on the horizontal axis. At the point 1.1 is the response at a frequency 10% higher than the design frequency; at the point 0.9 is the response at a frequency 10% lower than the design frequency.

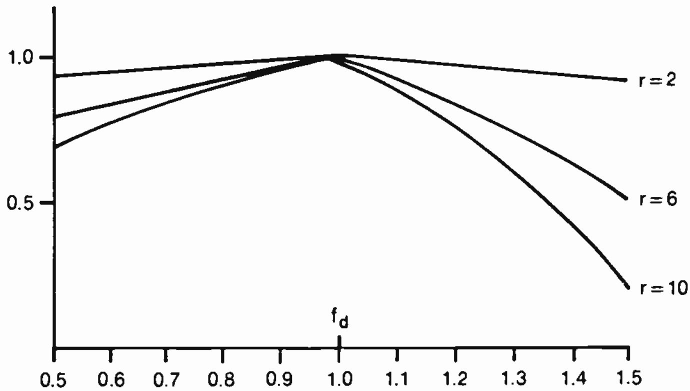


Fig. 10-11. Frequency response of L-networks.

Each curve is for a different value of impedance-transformation ratio. The curves show that the bandwidth of the network becomes smaller as the transformatin ratio becomes larger. Earlier we saw that circuit Q is given by

$$Q = \sqrt{\frac{R_p}{R_s} - 1}$$

where R_p / R_s is equal to the transformation ratio. This equation shows that a high impedance-transformation ratio corresponds to a high value of circuit Q, which we know means restricted bandwidth.

The curves of Fig. 10-11 show that the sidebands of a broadcast signal will be attenuated if the transformation ratio is greater than about 10. Substituting this into the equation in the preceding paragraph gives

$$Q = \sqrt{10 - 1} = \sqrt{9} = 3$$

As a general rule, a single network is never used to obtain an impedance-transformation ratio greater than 10:1, and the circuit Q is kept to 3 or less. Many problems in directional-antenna systems can be traced to using a single network for an impedance-transformation ratio greater than 10.

The curves of Fig. 10-11 can be made to apply to *L*-networks having a high-pass-filter configuration by simply letting the numbers on the horizontal axis represent the design frequency divided by the frequency of operation, instead of the inverse. Then 1.1 will correspond to a frequency 10% *below* the design frequency. In other words, the frequency scale will be reversed.

THREE-ELEMENT NETWORK

The most serious limitation of the 2-element *L*-network is that it introduces an amount of phase shift that depends on the impedance-transformation ratio. Once the transformation ratio is set, the amount of phase shift is set, and there is nothing we can do about it. By adding a third reactive element to our network, we can control the phase shift as well as the transformation ratio.

The 3-element network is widely used to perform the following functions:

1. Impedance transformation with arbitrary phase shift
2. Impedance transformation with specified phase shift
3. Control of phase shift without any impedance transformation.

The 3-element network can take either of the forms shown in Fig. 10-12. They are called *T* and pi networks because their

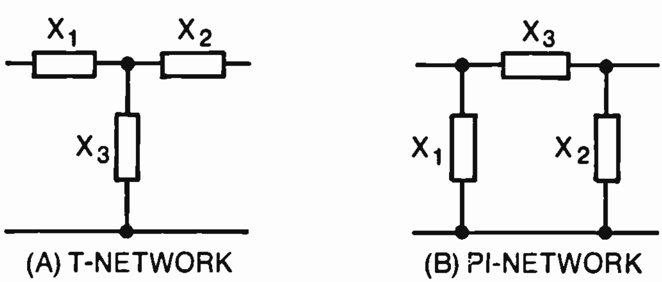


Fig. 10-12. T and Pi networks.

configurations resemble the English letter *T* and the Greek letter π. For any desired impedance transformation, we can use either type. In network theory we can show that for any given *T*-network there is an equivalent pi-network, and vice versa. As with the *L*-network, both inductive and capacitive elements must be used. That is, if two of the elements are inductive, the third must be capacitive.

Several combinations of reactive elements can be used to make up *T* or pi networks, some of which are shown in Fig. 10-13. The 3-element network most commonly used in broadcast work is the *T*-network in Fig. 10-13A. This configuration is popular because it has the configuration of a low-pass filter, and at broadcast frequencies and power levels, variable inductances are more practical than variable capacitances. The remainder of this discussion of 3-element networks is based on this configuration, but the principles of its operation apply to other possible configurations as well.

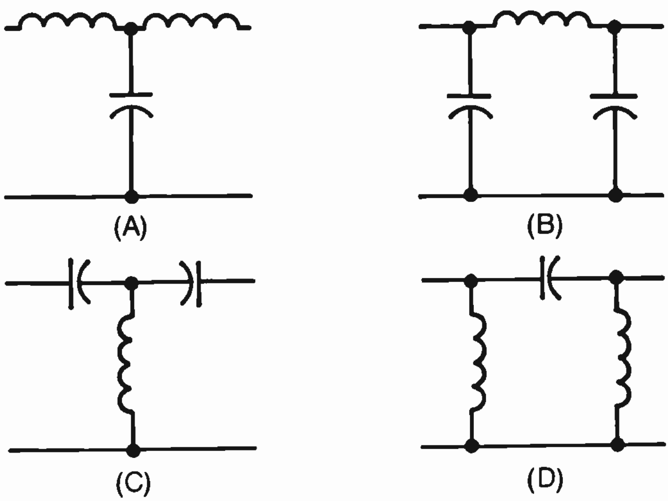


Fig. 10-13. Three-element networks.

In practical networks all three reactances are made variable. Variable capacitive reactance is obtained by connecting a variable inductor in series with a fixed capacitor as shown in Fig. 10-14.

Figure 10-15A shows two *L*-networks connected one after the other. By replacing the two capacitive reactances $X_3{}'$ and

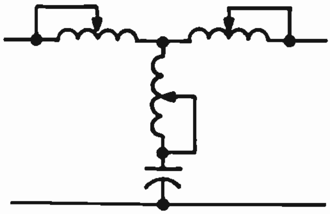


Fig. 10-14. Practical T-network commonly used in broadcast-antenna systems.

X_3'' with a single capacitor having a reactance of

$$X_3 = \frac{X_3' X_3''}{X_3' + X_3''}$$

we have the network of Fig. 10-15B. We can think of the *L*-section nearest the load as transforming the load impedance into some ficititious midpoint impedance Z_m , which is higher than both Z_L and Z_{in} . The second *L*-network transforms this fictitious impedance down to the desired value of input impedance Z_{in} .

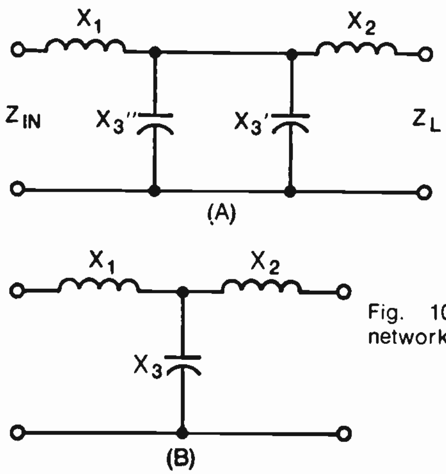


Fig. 10-15. Derivation of T-network.

If phase shift isn't important, we can design this network just as we would design two *L*-networks. We have one extra component value, which can be chosen arbitrarily.

Suppose we have the situation of Fig. 10-16, where we wish to match an 8-ohm resistive load to a 50-ohm line. We must make one arbitrary decision, so let's make the *Q* of the section closest to the load equal 2.5.

We can immediately find the value of X_2 in the circuit by simply multiplying Z_L by Q as shown in the figure. The midpoint impedance Z_m can be figured as the effective parallel resistance seen looking into the circuit. As shown, it works out to be 58 ohms. The capacitance X'_3 of the L-section closest to the load works out to be 23.3 ohms.

The rest of the problem consists of designing an L-network to transform the midpoint impedance of 58 ohms into the desired load impedance of 50 ohms. We first find the Q of this network, which we will call Q_1, to be 0.4. Now we can find that $X_1 = 20$ ohms and $X_3'' = 145$ ohms. By combining X_3' and X_3'' according to the rule of combining parallel reactances, we find that X_3 is also 20 ohms.

We can find the phase shift through this network by adding the phase shifts in the two L-sections.

$$\cos\gamma_1 = \frac{1}{\sqrt{r_1}} = \frac{1}{\sqrt{8/58}} = 0.37 \qquad \gamma_1 = 68.2^\circ$$

$$\cos\gamma_2 = \frac{1}{\sqrt{r_2}} = \frac{1}{\sqrt{58/50}} = 0.93 \qquad \gamma_2 = 21.8^\circ$$

$$\gamma = \gamma_1 + \gamma_2 = 68.2^\circ + 21.8^\circ = 90^\circ$$

This tecnhique can be used to analyze an existing T-network. The values of the elements of the network and the load impedance can be found with an impedance bridge, and the midpoint impedance can be computed. The phase shift through the network can then be calculated by adding the phase shifts of the two equivalent L-networks.

T-Network—90°

When all three elements of the T-network have the same numerical value, as in Fig. 10-16B, the network has some very interesting properties. In fact, it behaves very much like a section of transmission line that is 1/4 wavelength long. The common value of the three elements can be considered the *characteristic impedance* Z_0 of the network. The input and output impedances are related to the characteristic impedance of the network by the equation

$$Z_{in} = \frac{Z_0^{\ 2}}{Z_L}$$

where Z_L is the load impedance connected to one side of the network, and Z_{in} is the impedance that would be seen looking into the other end of the network. This is exactly the same as the equation relating the input and output impedances of a quarter-wave transmission line. As in the quarter-wave line, the output voltage and current lag the input voltage and current by 90°.

Another way of stating the relationship expressed by the above equation is to say that the characteristic impedance of the network is the *geometric mean* of the input and output impedances. The geometric mean of two numbers is the

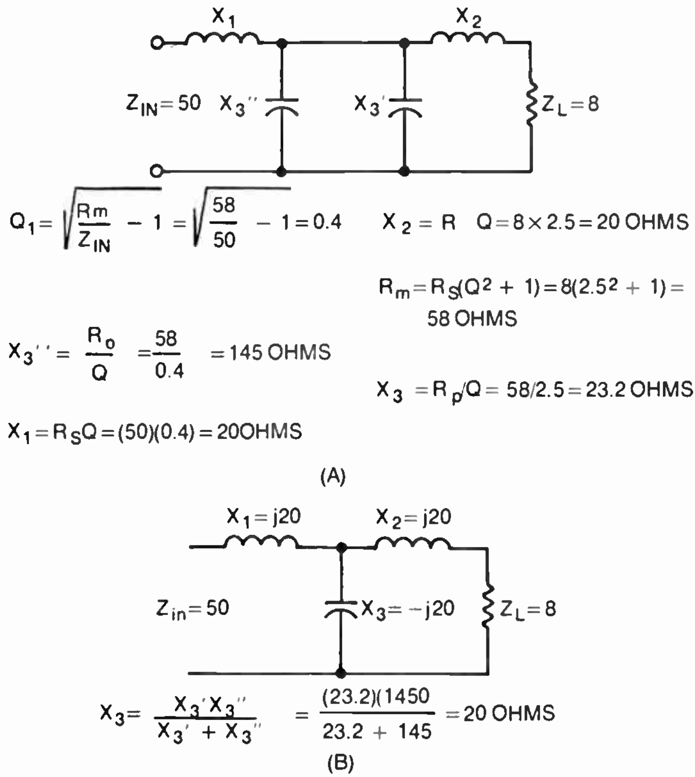


Fig. 10-16. Deriving a T-network from two L-networks.

number that would be halfway between them on the *C*-scale of a slide rule.

The 90° network is easy to design and adjust, because once the load and input impedances are known, the element values can be found easily.

For example, suppose that we wished to design a 90° network to match an 8-ohm load to a 50-ohm transmission line, as we did in Fig. 10-16. We can rearrange the above equation to solve for the characteristic impedance, giving

$$Z_0 = \sqrt{Z_{in}\ Z_L}$$

Substituting values for Z_{in} and Z_L gives

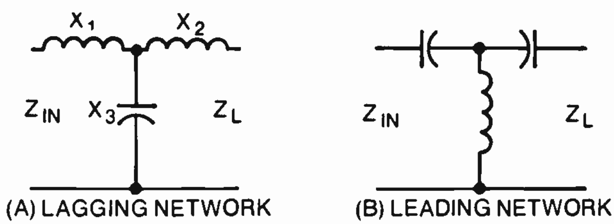
$$Z_0 = \sqrt{(50)(8)} = 20 \text{ ohms}$$

Thus all three elements of the network have a reactance of 20 ohms, which is exactly what we found in Fig. 10-16. Either X_1 and X_2 will be inductive and X_3 reactive, or else X_1 and X_2 will be capacitive and X_3 inductive (see Fig. 10-17). The network of Fig. 10-17A introduces a lagging phase shift of 90°, whereas the network of Fig. 10-17B introduces a leading phase shift of 90°.

The network of Fig. 10-17A can be seen to be the same network we designed in Fig. 10-16 by letting the *Q* of the *L*-section closest to the load equal 2.5. This example gives a little additional insight into the operation of the 90° network. In the next section we will see how the 90° *T*-network can be used as a phase shifter.

Phase Shifter—90°

A very common 90° *T*-network is the so-called 90° phase shifter of Fig. 10-18. In this *T*-network all three reactances as



$$X_1 = X_2 = X_3 = Z_0 = \sqrt{Z_{in} Z_L}$$

Fig. 10-17. Configurations of 90° T-networks.

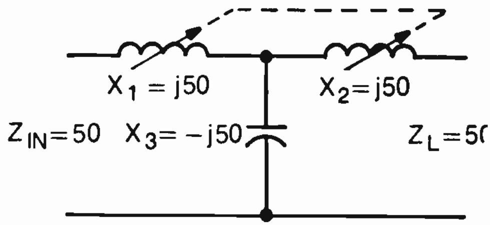


Fig. 10-18. A 90° phase shifter.

well as the input and output impedances have the same magnitude. Very frequently this value is 50 ohms and the network is inserted in a 50-ohm transmission line.

By varying X_1 and X_2, we can vary the phase shift of the network over a range of ±15° without disturbing the magnitude of the output current significantly. In the most common arrangement X_1 and X_2 are ganged to a single control so that they can be varied simultaneously. In some recent phase shifters only X_1 is variable; in most practical systems, this arrangement will disturb the magnitude of the load current more for a given amount of phase shift then when X_1 and X_2 are both variable.

The amount of phase shift, and the way the values of X_1 and X_2 affect it, depend on the load impedance seen looking back into the source. If the network were driven from a very high impedance or a constant-current source, X_1 would have no affect at all on the phase shift. If the network were driven from a very low impedance or a constant-voltage source, X_1 would have a large effect on the phase shift. In actual systems the source acts as neither a constant-voltage nor a constant-current source.

When the internal impedance of the source is equal to the characteristic impedance of the transmission line, varying X_1 alone produces a greater phase shift than when X_2 is also varied, but it also produces a greater change in load current. For a 15° phase shift the load current will change by about 4%.

When both X_1 and X_2 are varied, a greater change of reactance is required to produce a given phase shift, but the load current does not change nearly as much as when only X_1 is varied. For a 15° phase shift the load current will change by less than 1%.

It is difficult to analyze the behavior of a phase shifter in a practical system because its behavior is dependent on the impedance that will be seen looking back from the input terminals toward the transmitter. Although impedances are usually matched throughout the feeder system the impedance of the feeder system usually isn't matched to the internal impedance of the transmitter. The reason is simple: Although the condition of matched impedance represents maximum power transfer, it also represents the condition of 50% efficiency. If the impedance of a feeder system were matched to the internal impedance of the transmitter, half of the available power would be delivered to the antenna, and the other half would be dissipated in the transmitter. This is not a desirable mode of operation.

Figure 10-19 shows a transmitter connected through a line and a phase shifter to a load. The impedance is matched at the load and at the input and output of the phase shifter. The impedance is not matched at the input to the line leading from the transmitter. Most transmitters operate most efficiently when their load impedance is higher than their internal impedances. Inasmuch as the line is not matched at the sending end, the impedance seen looking back toward the transmitter from the input to the phase shifter depends on the length on the line. If the line is short compared to 1/4 wavelength at the operating frequency, the impedance seen looking back toward the transmitter will tend to be lower than the characteristic impedance of the line. The line will thus act something like a constant-voltage source. If, on the other hand, the line between the transmitter and phase shifter is about 1/4 wavelength, the impedance will be inverted and will be higher than the characteristic impedance of the line, and the line will

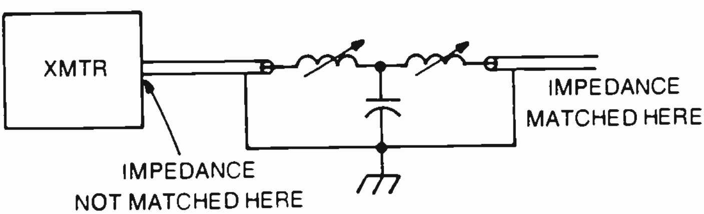


Fig. 10-19. Phase shifter connected between transmitter and load.

look more like a constant-current source. This uncertainty about the value of the source impedance feeding a network has no significant influence on an impedance-matching network, but it does have an influence on the way a phase-shifting network behaves.

GENERAL ANALYSIS AND DESIGN OF T-NETWORK

Network analysis and design are really two very different problems. The engineer trying to understand how the networks in a station operate analyzes the networks. That is, he starts with the values of the elements of the network, which he finds from a diagram or by measurements, and then determines the magnitudes and phases of the input and output currents or voltages. The designer, on the other hand, starts with the desired magnitudes and phases of the input and output currents, and then finds the values of the elements of a network that will satisfy these requirements.

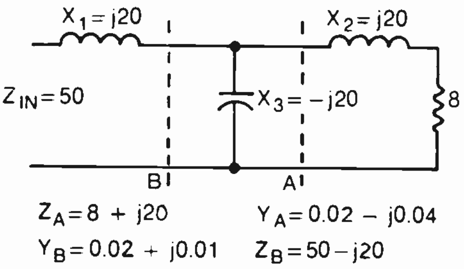


Fig. 10-20. Analysis of a T-network.

General T-Network Analysis

We will first consider the matter of analyzing networks. Fig. 10-21 shows a *T*-network for which the values of all of the reactances are given. The network is intended to match a load impedance of 8 ohms to the characteristic impedance of a 50-ohm line. As far as the solution of the network is concerned, it is a series–parallel network. So we have to combine impedances in series and parallel. This is normally a rather tedious procedure, but with a calculator that can handle vector calculations, there really isn't much work to it.

In analyzing any network it is easiest to start at the load and work back toward the source. In Fig. 10-20 we see that,

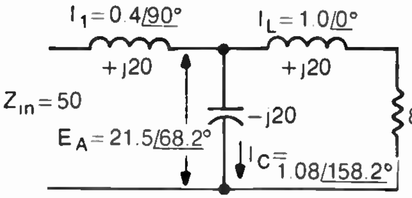


Fig. 10-21. Voltages and currents in a T-network.

looking at the load in series with X_2 —that is, looking to the right from line A, we have an impedance of

$$Z_A = 8 + j20 \text{ ohms}$$

This is an admittance of

$$Y_A = 0.02 - j0.04 \text{ mho}$$

The next element of the network is a capacitive reactance of $-j20$ ohms. This is the same as a capacitive susceptance of $j0.05$ mho. Since this element is in parallel with the other admittance, we can add it algebraically to the susceptance of the load; that is,

$$Y_B = 0.02 - j0.04 + j0.05 = 0.02 + j0.01 \text{ mho}$$

To consider the last series element, it is convenient to convert back to impedance.

$$Z_B = 50 - j20 \text{ ohms}$$

Adding the reactance of $j50$ gives us

$$Z = 50 - j20 + j20 = 50 \text{ ohms}$$

Now that we know the value of admittance or impedance at each point in the network we can rather easily find the phases of the currents. The easy way to do this is to assume some value of load current, as in Fig. 10-21. One ampere is as good a choice as any. This current flows through X_2 and the load, so we can figure the voltage at the midpoint of the network by Ohm's law.

$$E_A = I(R + jX_2) = 1(8 + j20) = 21.5 \angle 68.2^\circ \text{V}$$

We can use Ohm's law again to find the current through the capacitor.

$$I_C = \frac{E}{X} = \frac{21.5 \angle 68.2^\circ}{20 \angle -90^\circ} = 1.08 \angle 158.2^\circ \text{A}$$

The current lags the voltage across the capacitor by 90°, which is exactly what it should do. We now know the two currents I_L and I_C, so we can find the current through I_1, which is also the input current of the network, by using Kirchhoff's current law.

$$I_1 = I_L + I_C = 1 \angle 0° + 1.08 \angle 158.2° = 0.4 \angle 90°$$

We now have all of the currents in the network. The ratio of the output current to the input current is

$$\frac{I_L}{I_1} = \frac{1.0 \angle 0°}{0.4 \angle 90°} = 2.5 \angle -90°$$

Inasmuch powers are constant, the impedance-transformation ratio is the square of the current ratio. The phase shift through the network is −90°

Any network can be analyzed by using these procedures. When a calculator is available, all of the computations can be made in a few minutes.

DESIGN OF T-NETWORK

By contrast, the *design* procedure is tedious and at times confusing. If all of the element values in the network are determined by a detailed analysis, the mathematical manipulations can be tedious indeed. Fortunately, there are formulas that can be used to design a *T*-network for any desired impedance transformation and phase shift. Of course, the consideration of keeping the impedance-transformation ratio to 10:1 or less applies in the *T*-network just as it does in the *L*-network. The formulas for the value of each of the elements of a *T*-network are as follows:

$$X_1 = \frac{\sqrt{R_{in} R_L}}{\sin \gamma} - \frac{R_{in}}{\tan \gamma}$$

$$X_2 = \frac{R_{in} \; R_L}{\sin \gamma} - \frac{R_L}{\tan \gamma}$$

$$X_3 = -\frac{\sqrt{R_{in} \; R_L}}{\sin \gamma}$$

where R_{in} = desired input resistance
R_L = load resistance
γ = desired phase angle between input and output currents

To understand some of the implications of the above formulas, let us go back for a moment to the *L*-network. In considering an *L*-network, we found that the phase shift is given by

$$\cos\gamma = \frac{1}{\sqrt{r}}$$

Now, if we use the above design formulas for a *T*-network having a phase shift given by this equation, we find that the value for either X_1 or X_2 becomes zero. This means that we won't have any inductor at all for X_1 or X_2, and the network will become an *L*-network. This shows that the only reason we need three elements is to get some value of phase shift other than what we would get with an *L*-network for the same impedance-transformation ratio.

Another interesting implication of the formulas is that if we wished to get a phase shift between zero and the value we would obtain from an *L*-network, X_1 would be negative, meaning that we would need a capacitive reactance at this point instead of the inductive reactance that we have shown. That is, if

$$0 < \gamma < \cos^{-1}\left(\frac{1}{\sqrt{r}}\right)$$

then X_1 will be a capacitor.

Based on our formulas for the values of the *T*-network, we can define four different types of *T*-networks, each of which provides a phase shift in a given range, as shown in Fig. 10-22.

Whenever an impedance-transforming or phase-shifting network is designed, it should be analyzed in detail to determine the relative magnitudes of the currents, the bandwidth, and the efficiency.

Bandwidth of T-Network

Like the *L*-network, the *T*-network is designed on the assumption that the inductive and capacitive reactances have

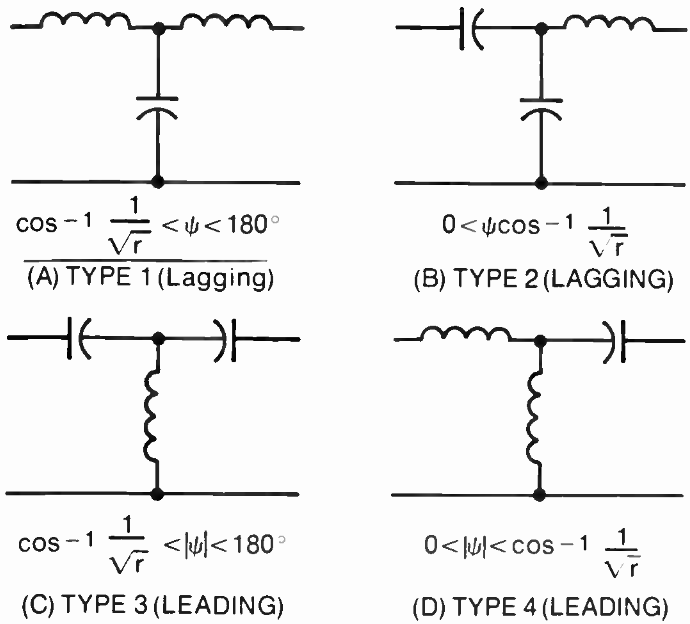


Fig. 10-22. The various types of T-networks.

constant values. Inasmuch as the reactance of a coil or capacitor changes as frequency changes, we can expect ideal behavior from the network at only a single frequency. At other frequencies the response will be somewhat different.

Figure 10-23 shows how output current varies with frequency. Note that as the impedance-transformation ratio r becomes greater, the effect of frequency is more pronounced. For this reason, in the *T*-network (as in the *L*-network), it is

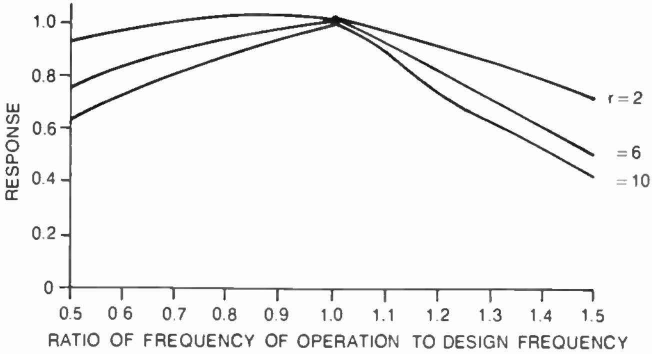


Fig. 10-23. Bandwidth of T-network.

not advisable to attempt an impedance transformation greater than 10:1 in a single network. Sometimes, when a greater impedance-transformation ratio is required, a *T*-network and an *L*-network are used in tandem.

NETWORK EFFICIENCY

If a network is inefficient, not only will power that could have been radiated be lost, but the components of the network will dissipate the wasted energy in the form of heat. This means that the components will require higher ratings than otherwise, and inasmuch as they will operate at higher temperatures, they will be more apt to fail or change value.

For all practical purposes, we can make a couple of assumptions that make calculating the power lost in a network a simple matter.

1. The currents in the network can be calculated by considering the elements of the network to be pure reactances with no losses.
2. The losses in capacitors can be neglected because the losses in the inductors are so much greater.

Based on these assumptions, it is rather easy to compute the loss in a network. When a network is properly adjusted, the input and load impedances will be pure resistances. The power in or out of the network will therefore be

$$P_l = I_{in}^2 R_1 + I_l^2 R_2$$

Since we are assuming that all of the losses take place in the resistance in the inductances in the circuit, the network losses will be

$$P_l = |I_{in}|^2 R_1 + |I_l|^2 R$$

where R_1 & R_2 are the resistances associated with X_1 and X_2. These resistances are shown in Fig. 10-24 and may be determined from measurements with an impedance bridge.

By a lengthy mathematical analysis of the loss in networks, we could show that the efficiency of a network depends only on the *Q* of the inductances, the impedance-transformation ratio, and the amount of phase shift. The

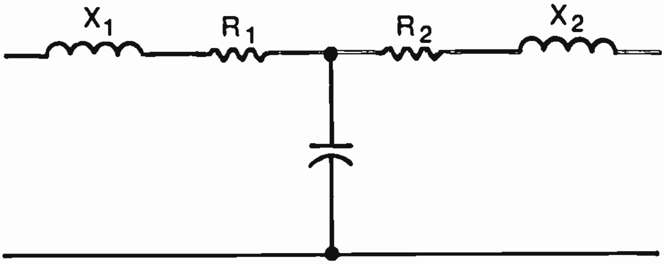


Fig. 10-24. Loss resistances in a T-network.

greater the impedance-transformation ratio, the lower the efficiency of the network. And the more the phase shift differs from 90° (for the type 1 network), the lower the efficiency.

The analysis would disclose a rather surprising fact; that is, the efficiency of the network is not dependent on the number of coils used. Thus, the efficiency of a matching network can sometimes be improved by using two networks in tandem instead of one network.

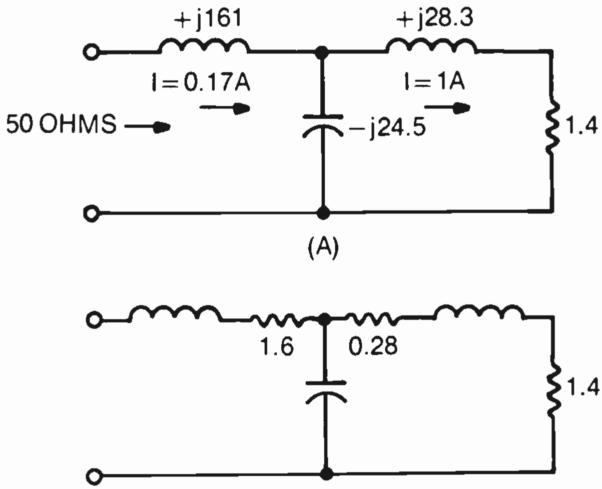
The efficiency of an *L*-network can be determined by the simple relationship

$$\text{Efficiency} = \frac{R_L}{R_L \times R_s}$$

where R_L is the load resistance, and R_s is the series loss resistance of the network. Inasmuch as a *T*-network is simply two *L*-networks in tendem, the efficiency is simply the product of the efficiencies of the two *L*-networks.

Figure 10-25 shows a *T*-network designed to produce a phase shift of 160° and an impedance-transformation ratio of 36. The coils are assumed to each have a *Q* of 100. The currents, based on a one-ampere load current, are shown in the figure, as is the efficiency. This network, which violates our fundamental rule that the impedance-transformation ratio should not be greater than about 10, is a good example of why we try to keep the impedance-transformation ratio low. The efficiency of this network is only about 81%. This means that 19% of the power fed to the network is dissipated in the network and does not reach the antenna.

In Fig. 10-26 we have two *T*-networks in tandem. Each has an impedance-transformation ratio of 6 and a phase shift of



$$P_l = 0.17^2(1.6) + 1^2(0.28)$$

$$= 0.33\text{W}$$

$$P_{out} = 1^2(1.4) = 1.4\text{W}$$

$$\text{Efficiency, \%} = \frac{P_{out}}{P_{out} + P_l} = 1.4/1.73 = 81\%$$

Fig. 10-25. A 36:1 T-network.

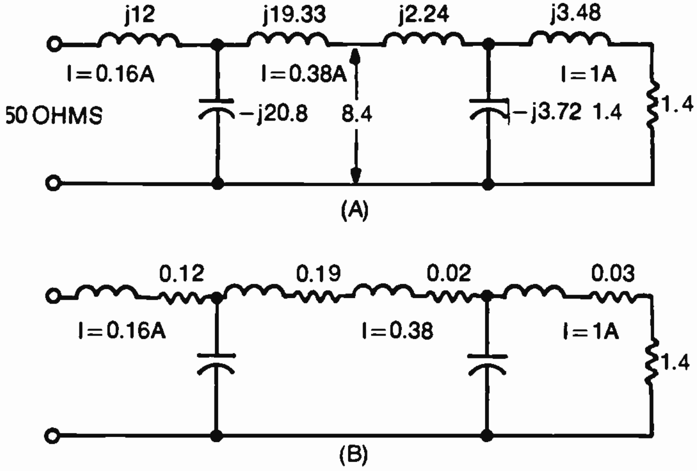
80°. The impedance-transformation ratio of the two networks is equal to the product of their individual impedance-transformation ratios—in this case, 36. The phase shift through tandem networks is equal to the algebraic sum of their individual phase shifts—in this case, 160°. Thus the two tandem networks of Fig. 10-26 do exactly what the one network of Fig. 10-25 does. As shown in Fig. 10-26, however, the efficiency of the tandem networks is about 96%.

This example shows that two networks using a total of four coils, each having Q of 100, are more efficient than a single network using only two coils each having a Q of 100. Whenever an engineer finds that in spite of all his efforts he can't seem to get an antenna feeder system to operate efficiently, he should check the efficiency to see if it might be improved considerably by using two networks rather than one.

One conclusion to be drawn from this discussion of efficiency is that it is very difficult to achieve high efficiencies with load impedances that are very low compared to the characteristic impedance of the transmission line. For this reason, designers try to avoid using towers that have very low driving-point impedances. Of course, if the power fed to such a tower is very small, a lower efficiency can be tolerated. When a high efficiency is required, two networks may have to be used in tandem.

When we consider the efficiency in matching a transmission line to an antenna, we can't ignore the characteristics of the antenna itself. When an inductance is used to tune out the capacitive reactance of a short antenna, the efficiency of the combination is given by

$$\text{Efficiency, \%} = \frac{Q_L}{Q_L + Q_A}$$



$P_4 = 0.16^2(0.12) + 0.38^2(0.19 + 0.02) + 1(0.03)$

$P_R = 0.063W$

$P_{out} = 1^2(1.4) = 1.4W$

$$\text{Efficiency, \%} = \frac{P_{out}}{P_{out} + P_l} = 96\%$$

Fig. 10-26. Improved efficiency with two T-networks in tandem.

where Q_A = Q of antenna (base reactance divided by base resistance) Q_L = Q of inductance

Obviously, from this equation, the Q of the inductance should be as high as possible, and the Q of the antenna should be as low as possible. This is one reason why standard broadcast stations rarely use antennas much shorter than 90°. An antenna 50° high would have a Q of about 100. If the coil used for matching (say, in an L-network) had a Q of 100, the efficiency of the combination would only be about 50%.

SERIES RESONANCE

The familiar series-resonant circuit is very frequently used in directional-antenna feeder systems. It has the following applications, each of which will be discussed in the following paragraphs:

1. It can give a phase shift of a few degrees or either side of 0°.
2. It provides a means of obtaining a variable capacitive reactance by using a variable inductance.
3. It provides increased harmonic reduction in a T-network.

Figure 10-27 shows the series-resonant circuit together with a plot of its impedance and phase shift. The impedance is seen to be very low at resonance, and higher on either side of

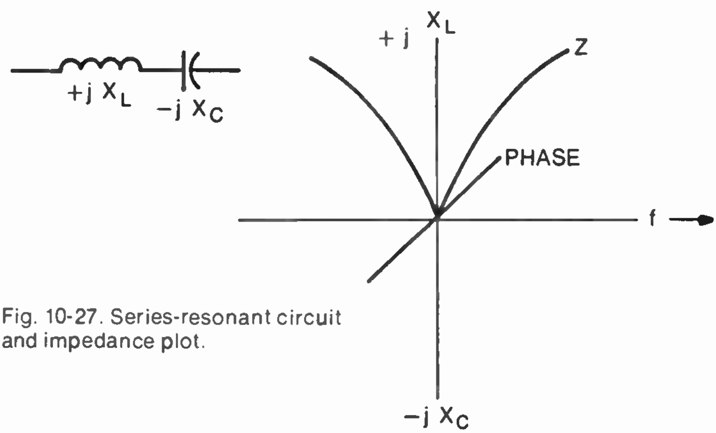


Fig. 10-27. Series-resonant circuit and impedance plot.

resonance. The equation for the magnitude of the impedance of a series-resonant circuit is

$$Z = R + j(X_L - X_C)$$

and the phase shift θ between the voltage and current is

$$\theta = \tan^{-1} \frac{X_L - X_C}{R}$$

The Q of the circuit is defined as X_L/R, where R includes not only the intrinsic resistance of the inductance but any other resistance in series with the circuit.

An important consideration in the use of series-resonant circuits in broadcasting is the voltage rise across the inductance and capacitance. The voltages across the capacitance and inductance are opposite in the phase, thus they cancel out as far as the external circuit is concerned. Nevertheless, high reactive voltages actually appear across the inductance and capacitance. They are approximately equal to the voltage applied to the circuit, multiplied by the Q of the circuit. If a circuit has a Q of 100 and one volt is applied, the voltage across the inductance and capacitance is in each case about 100V. At the voltages encountered in broadcasting, thc voltage rise would be prohibitive unless the Q were held to a very low value. As with the L-network, the Q of the circuit should be no greater than 3.

Zero-Degree Phase Shifter

Figure 10-28 shows a series-resonant circuit used as a phase shifter. It is called a *zero-degree* phase shifter because it shifts the phase up to about ±15° around 0°, as compared to the T-network phase shifter, which shifts the phase about 90°. The 0° phase shifter is used where the characteristics of a feeder system are such that a 90° phase shift could not be tolerated.

As shown in Fig. 10-28B, the complete circuit includes not only the resistance of the load, (R_L) but also the resistance of the source (R_s). The amount by which the phase shifts for a given change in reactance depends on the Q of the circuit hence upon the resistance of the source. We assume that the load resistance has been properly transformed by networks so

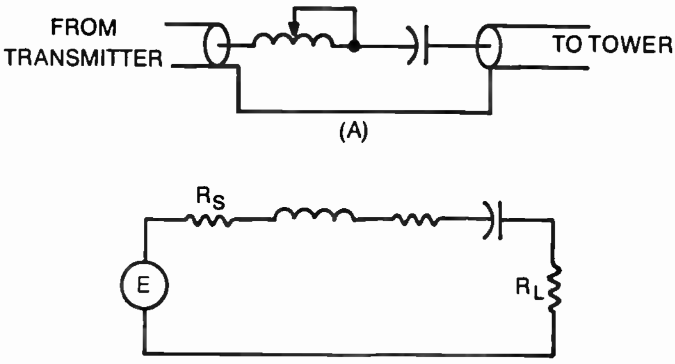


Fig. 10-28. Zero-degree phase shifter.

that it looks like a pure resistance equal to the characteristic impedance of the line.

The Q of the complete circuit must be kept low so as not to unduly restrict the bandwidth and so as to permit phase shifting without affecting the load current.

Obtaining Variable Capacitive Reactance

At the power levels used in standard broadcasting, variable capacitors are large and expensive, although vacuum-type variable capacitors are becoming practical. It is customary in broadcasting to use variable inductors rather than variable capacitors. Hence, when we need a variable capacitive reactance, we use the arrangement of Fig. 1-29, which is really just a series-resonant circuit used on the low side of the resonant frequency.

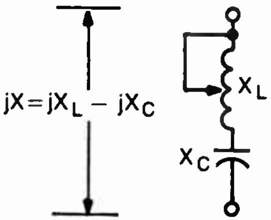


Fig. 10-29. Obtaining variable capacitor reactance.

This circuit is not without its pitfalls. Suppose, for example, that we need a variable capacitive reactance of approximately 50 ohms at an operating frequency of 1 MHz. We could get this reactance with the arrangement of Fig.

10-30A which calls for a capacitor having a reactance of 100 ohms. At 1 MHz this would mean a capacitance of 0.0015 μF. The calculations in the figure show the effect of a 1% change in the value of either inductive or capacitive reactance, such as might result from temperature changes. Note that if one of the series reactances changes value by 1%, the net reactance of the circuit will change by 2%. This is not objectionable.

Suppose, however, that the designer tried to cut costs by using the circuit of Fig. 10-30B. Here the value of the capacitor is only 160 pF. At the power rating required for broadcasting, this would represent a substantial saving. However, as shown in the figure, if either of the series reactances changed value by 1%, the net reactance of the circuit would change by 20%. Such an arrangement would result in a very unstable system. Unfortunately, such arrangements are occasionally found in a broadcast-antenna system, sometimes as a result of building a network from whatever components happen to be available.

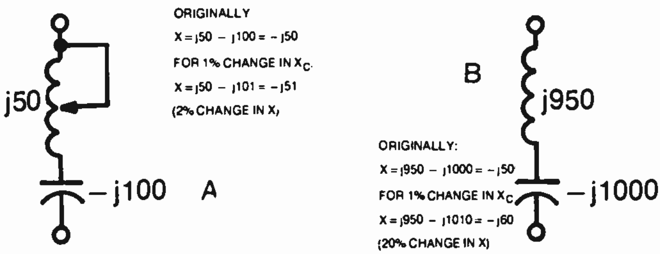


Fig. 10-30. Stability of series LC circuit.

Harmonic Reduction

Figure 10-31 shows a *T*-network that might be used for impedance matching in a broadcast-antenna system. The difference between this network and the basic *T*-network is that there is both a capacitor and an inductor in the shunt arm of this *T*-network. As we saw, an *LC* combination can be used to provide a variable capacitive reactance, using a variable inductor. However, it can be used for another purpose.

For purposes of impedance matching, the shunt arm must have some specified value of capacitive reactance, say, 20 ohms. Hence, in this case, the fixed capacitor and variable inductor are selected so that the difference between their reactances will be 20 ohms. This means that the circuit will be operating at a frequency below the resonant frequency. Thus the impedance will be minimum at some higher frequency, where the circuit is resonant. By proper choice of the values of the capacitance and inductance, we can make the resonant frequency occur at the second, or any other, harmonic of the carrier frequency. This arrangement is particularly helpful in reducing second-harmonic radiation. The desired values can be selected from the equations below.

At design frequency f_d

$$2\pi f_d L - \frac{1}{2\pi f_d C} = X_3$$

At harmonic frequency f_h

$$2\pi f_h L - \frac{1}{2\pi f_h C} = 0$$

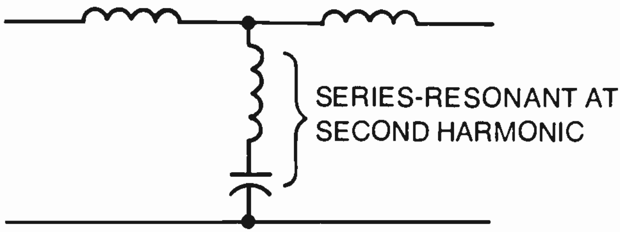


Fig. 10-31. A T-network used for harmonic rejection.

MATCHING COMPLEX LOAD IMPEDANCES

So far, we have seen that within certain practical limits, we can transform any value of load resistance into some other value of resistance, using an *L*- or *T*-network. In practice, the driving-point impedance of an antenna tower is rarely a pure resistance. There is almost always a reactive component. To obtain maximum power transfer and minimum reflection, this complex impedance must be transformed into a pure resistance equal to the characteristic impedance of the transmission line. This is accomplished by inserting the opposite type of impedance in series with the lead to the base of the tower. In many cases, this additional reactance can be a part of the matching network.

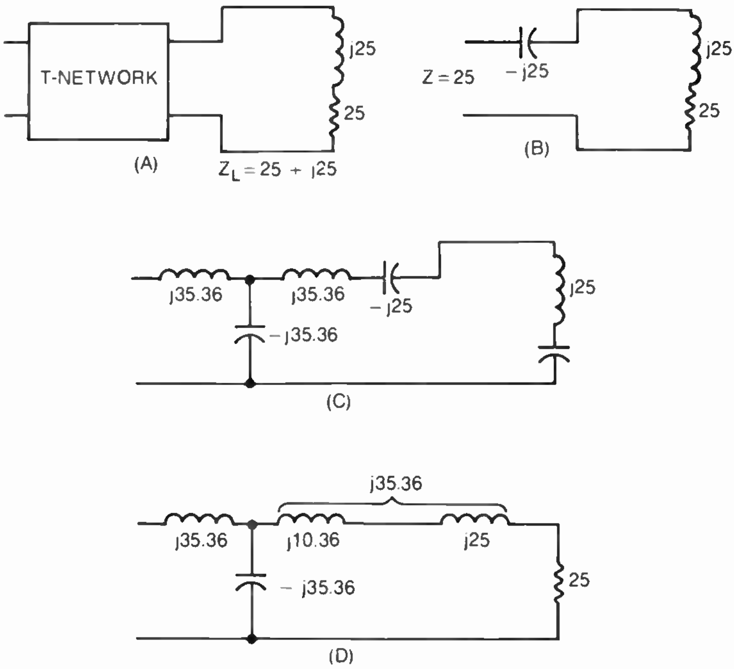


Fig. 10-32. Handling reactive loads with a T-network.

Figure 10-32A shows a load impedance of 25 + *j*25 ohms. If we add a capacitive reactance of 25 ohms in series with this load, as shown in Fig. 10-32B, it will then look like a pure resistance of 25 ohms. We can then design an *L*-network or *T*-network to transform the 25 ohms into 50 ohms to match a transmission line. This is shown in Fig. 10-32C, where we have a 90° *T*-network in which the value of each of the elements is 35.36 ohms—the geometric mean of the load and driving-point impedances.

It isn't necessary to include the series capacitor of Fig. 10-32C. We can instead let the 25-ohm reactance of the load be part of the network, as shown in Fig. 10-32D. This means that the reactance of the network element only has to be 10.36 ohms, the difference between the 25-ohm reactance in the load impedance and the desired value of 35.36 ohms.

This technique can be used to handle any practical value of load impedance. Remember that we have several configurations of *L*- and *T*-networks to choose from.

Chapter 11

Feeder Systems for Standard Broadcast Antennas

Before a tower can radiate a signal, it must receive a signal from the transmitter. For proper operation it is essential that the line carrying the energy to the tower does not radiate energy. All of the radiation must be done by the tower. The distribution of energy to the towers is accomplished by an assembly of components called the *feeder system*.

In case of a nondirectional antenna, the feeder system consists of a transmission line together with a network that matches the impedance seen at the base of the tower to the characteristic impedance of the transmission line. The feeder system of a directional antenna performs many more functions, including the following:

1. Controlling the magnitude of the current fed to each tower
2. Controlling the relative phase of the current fed to each tower
3. Transforming the impedance seen at the common point of the system to a value suitable for properly loading the transmitter.

The feeder system also includes components for supplying current to the tower lights, providing lightning protection, and measuring the amplitudes and phases of the tower currents.

These three subjects are covered in later chapters. Since the nondirectional-antenna feeder system is really a very greatly simplified version of the more general directional-antenna system, it will not be covered specifically.

Once a directional antenna has been installed, the only adjustments for controlling the operating parameters of the system are located in the feeder system. In fact, just about all of the work in the day-to-day operation of a directional-antenna system consists of reading instruments and adjusting or tuning controls in the feeder system.

FEEDING THE TOWER

The most common way of feeding RF energy to a tower is to apply the signal across a base insulator that is placed between the bottom of the tower and the ground system (Fig. 11-1). There are disadvantages to this approach, however. When a tower is insulated from ground, there is the added cost of the base insulator, which msut support the weight of the tower, and there are difficulties in feeding power to tower lights and providing lightning protection. Hence many alternate methods of feeding energy to towers have been proposed through the years.

The only alternate system that has gained even limited acceptance is the *shunt* arrangement shown in Fig. 11-1B. Here the base of the tower is connected directly to ground, and the energy is fed to the tower through a slanted wire, as shown. The system may be viewed as a single-turn, 3-sided loop that consists of the slant wire, the bottom section of the tower, and the ground path. The magnetic field from this loop induces a voltage in the tower, thus coupling energy to it.

In theory it is possible to select a connection point on the tower for the slant wire so that the resistive component of the impedance will be equal to the characteristic impedance of the transmission line. The reactive component, which will always be inductive, can be tuned out by the capacitor connected in series with the slant wire. Care must be taken to minimize the losses in the ground path. Usually a copper strap is connected between the outer conductor of the transmission line and the center of the ground system.

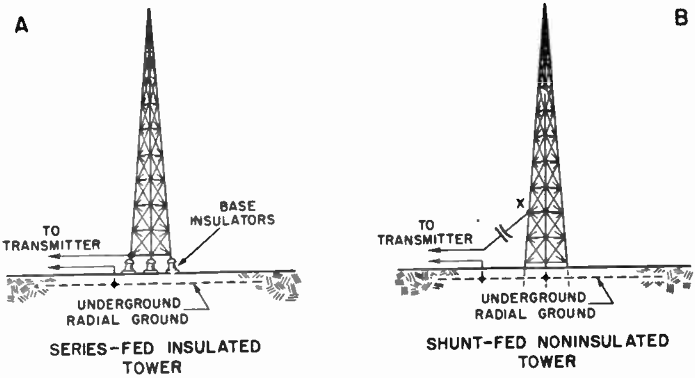


Fig. 11-1. Series- and shunt-fed tower radiators.

Although limited success has been achieved with shunt feeding of single towers, the method is seldom used in directional-antenna arrays. One reason is that the proper point for connecting the slant wire is not easy to find theoretically, so that a great deal of cut-and-dry work is required for optimum coupling. Also, radiation from the slant wire can be troublesome.

With the *series* arrangement (Fig. 11-1A), the two principal considerations are the driving-point impedance of the tower and the voltage across the base insulator. With a tower 90° in height, the voltage across the base insulator is minimum, being in the order of 200V to 250V for 1 kW of radiated power. The voltage across the base insulator is greatest when the antenna height is about 180°. The base voltage can then be as high as 1900V for 1 kW of radiated power. Since voltage is proportional to the square root of power, the base voltages for other powers can be obtained by multiplying the above approximate figures by the square root of the radiated power in kilowatts.

LAYOUT OF FEEDER SYSTEM

The geometrical layout of a particular feeder system depends on the type of antenna array and the location of the transmitter building. Each transmission line must terminate at a tower, and the lines must converge at the point where the power-dividing and phase-shifting equipment is located. Figure 11-2 shows several possible layouts of feeder systems.

There is little that the broadcast engineer can do about the layout of the system. But, if it has serious limitations, he can make changes whenever a major modification is initiated. In general, there are three rules of thumb for making a good layout.

1. The phasor equipment—that is, the power-dividing and phase-shifting networks—should all be located in one place. There is always some interaction between the power-division and phase-shift controls, and if they are separated, adjustment becomes unnecessarily complicated.
2. The transmission lines should not be any longer than necessary. Keeping the lines short minimizes losses in the system.
3. When practicable, the phasor should be located in the transmitter building, bringing most of the adjustments to one central location.

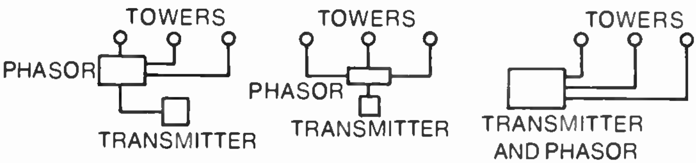


Fig. 11-2. Possible feeder-system layouts.

TYPICAL FEEDER SYSTEM

Figure 11-3 shows a block diagram of a feeder system for a 2-tower directional-antenna array. Tower 1 is the reference tower of the array. To understand how the feeder system operates, we will start at the reference tower and trace the signal back to the transmitter. Then we will do the same thing with tower 2. For convenience, we will consider the phase angle of the current entering tower 1 to be zero. The magnitude of the current flowing into this tower is 3.65A.

The first element to be considered is an impedance-transforming network (in the antenna-coupler block) at the base of tower 1. This network transforms the driving-point impedance of the tower in to the characteristic impedance of the line. In this particular case, the driving-point impedance of

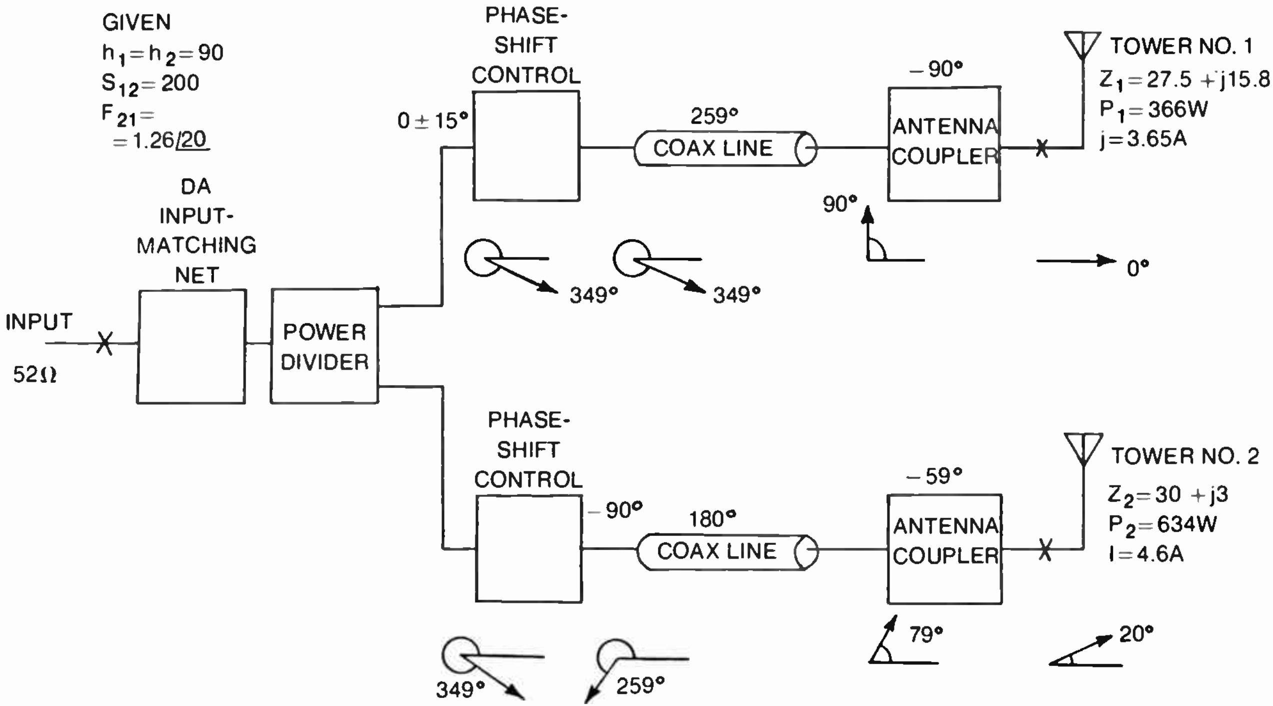


Fig. 11-3. Block diagram of directional-antenna feeder system.

tower 1 is 27.5 + $j16$ ohms, and the characteristic impedance of the transmisson line is 50 ohms. Either an *L* or *T*-network might be used to accomplish the desired impedance transformation; usually the choice is based on phase-shift considerations. If a *T*-network is used, the phase shift can be adjusted to any desired value. If an *L*-network is used, we have to live with whatever phase shift the particular impedance-transformation produces. In the system of Fig. 11-3 a 90° *T*-network is assumed. Its 90° phase shift, which is lagging, must be taken into consideration in figuring the relative phase of the currents feeding the two towers. In the figure the current feeding tower 1 is represented by a vector at 0°. We must, therefore, represent the current at the input of the matching network by a vector that is lagging the reference vector by 90° as shown.

The electrical length of the transmission line from the network at tower 1 back to the phasing equipment (phasor) is 259°. Inasmuch as the signal is delayed in passing through the transmission line, this phase angle is also lagging. Thus the current at the input to the transmission line is represented by a vector that lags the reference by 90° + 259°, or 349°.

At this point in the particular system we are studying, there is a 0° phase shifter. In many systems there is no phase adjustment in the line to the reference tower. One phase shift adjustment is all that is required in a 2-tower array, as we are interested in the phase between the currents to the two towers, and changing the phase of either current will change the phase angle between them. The phase shifter in this line, although it isn't strictly necessary, adds flexibility to the system and provides a wider range of adjustment of the phase angle between the tower currents. We saw in an earlier chapter that a phase shifter can provide about 15° of phase shift without seriously affecting the magnitude of the current. If a phase shifter is provided in each line of a 2-tower array, we can shift the phase angle between the currents by as much as 30° without seriously disturbing the magnitudes of the currents significantly.

Before any adjustments are made, the 0° phase shifter doesn't introduce any phase shift. Hence we can represent the current at the input to the phase shifter by the same vector that we used to represent the current at its output—that is, a phasor with a lagging angle of 349° with respect to our reference (the current to tower 1). At this point the line connects to a network called a *power divider* which we will look into later in this chapter. Now, let's look at the phase shift in the line feeding tower 2.

At the base of tower 2 we have the same problem of transforming a complex driving-point impedance into the characteristic impedance of the transmission line. In this case, the driving-point impedance is $30 + j3$ ohms, and the characteristic impedance of the line is 50 ohms. The magnitude of the current feeding tower 2 is 4.6A, and this current must load the current of tower 1 by 20°. Thus we can represent the current of tower 2 by a vector drawn at an angle of 20° with respect to our reference.

Again, either a *T* or *L*-network may be used to accomplish the required impedance transformation. In this case, a *T*-network is used, and its phase shift is adjusted to −59°. The reason for this particular value of phase shift will become apparent as we proceed.

The current at the input to the impedance-transforming network at the base of tower 2 can be represented by a vector with a lagging angle of 20° + 59°, or 79°. The length of the transmission line from tower 2 back to the phaser is 180°, so the input current to the line can be represented by a vector with a lagging angle of 79° + 180°, or 259°.

At this point a 90° phase shifter is inserted, and the current at the input to the phase shifter can be represented by a vector with a lagging angle of 259° + 90°, or 349°. We can see that the required phase of current from the power divider to the line feeding tower 2 is exactly the same as that for the line feeding tower 1. In other words, the two currents coming from the power divider are in phase, and by proper selection of components in the feeder system, the currents at the towers will have the required phase difference of 20°.

It is important for any broadcast engineer in charge of a station with a directional antenna to know how the phase shifts in the feeder system combine to produce the required phase angles between tower currents.

It is interesting to note that the transmission line feeding tower 2 in our example is exactly 1/2 wavelength long. This would happen because of the geographical spacing between the towers and the transmitter house, not for any electrical reason. In a half-wave line the driving-point impedance is exactly equal to the load impedance. If the load impedance has any value other than the characteristic impedance of the line, there will be a standing wave on the line, but it's presence may go unnoticed. Sometimes this particular length of transmission line appears easy to match—until a standing wave causes the line to fail.

POWER DIVIDER

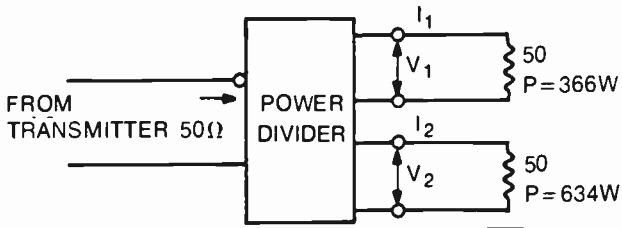
The purpose of a power divider is to take power from a single transmission line from a transmitter and distribute it to several transmission lines in such a way that the proper magnitude of current is supplied to each tower in an antenna array. The problem is shown schematically in Fig. 11-4A. Here we have two 50-ohm loads, representing the transmission lines to towers 1 and 2 in the array of Fig 11-3. As specified in that figure, 366W must be delivered to load 1, and 634W to load 2. This is to be accomplished by a yet unspecified circuit that has a driving-point impedance of 50 ohms. Applying Ohm's law to the problem, we can calculate the voltage and current applied to each load that will result in the proper amount of power being delivered.

Low-Frequency Power Divider

In ordinary 60 Hz power systems we have many power-division problems—delivering the proper amount of power to each lamp in a home, for example. At power-line frequencies power division is accomplished by controlling the amount of resistance of the load and operating the system on a constant-voltage basis. A low-resistance lamp draws more power from the line than a high-resistance lamp. In this type of system, impedances are intentionally unmatched. We don't want a load that will draw the maximum possible amount of power from a generating station!

It is possible, however, to design a low-frequency power-dividing system that operates on a constant impedance level. In fact, it is a good idea for us to do this in order to get a better idea of power division between impedances of the same value.

Referring to Fig. 11-4, we see that load 1 (50 ohms) will require a voltage of 135.5V and a current of 2.71A to draw 366W. Similarly, load 2 (also 50 ohms) will require a voltage of 178V and a current of 3.56A to draw 634W. Also, the 50-ohm driving-point impedance of whatever is in the box of Fig. 11-4 will require a voltage of 223.6V and a current of 4.47A to draw the total required power of 1000W for the two loads (366 + 634 = 1000W).



$$I_1 = \sqrt{\frac{P_1}{50}} = \sqrt{\frac{366}{50}} = 2.71\text{A} \qquad I_2 = \sqrt{\frac{P_2}{50}} = \sqrt{\frac{634}{50}} = 3.56\text{A}$$

$$V_1 = \sqrt{P_1R} = \sqrt{360 \times 50} = 135.3\text{V} \qquad E_2 = \sqrt{P_2R} = \sqrt{634 \times 50} = 178\text{V}$$

Fig. 11-4. Power divider problem.

We can easily specify a transformer to accomplish the necessary power division. Working on a voltage basis, we can specify a transformer like that shown in Fig. 11-5A, which has a turns ratio such that when 223.6V is applied to the primary, the proper voltages will appear at each secondary. The ratios work out to be 0.61 for load 1 and 0.80 for load 2. That is, for every 100 turns on the primary, there will be 61 turns on secondary 1 and 80 turns on secondary 2. (You should carefully calculate the voltage, current, and impedance values at each point in the circuit of Fig. 11-5A to get a good feeling for what is going on. This will be helpful in understanding how an RF power divider operates.) In Fig. 11-5B we have simply replaced a transformer having a primary and two secondaries with an *autotransformer* having two taps on its single winding. A similar arrangement is used for RF power division.

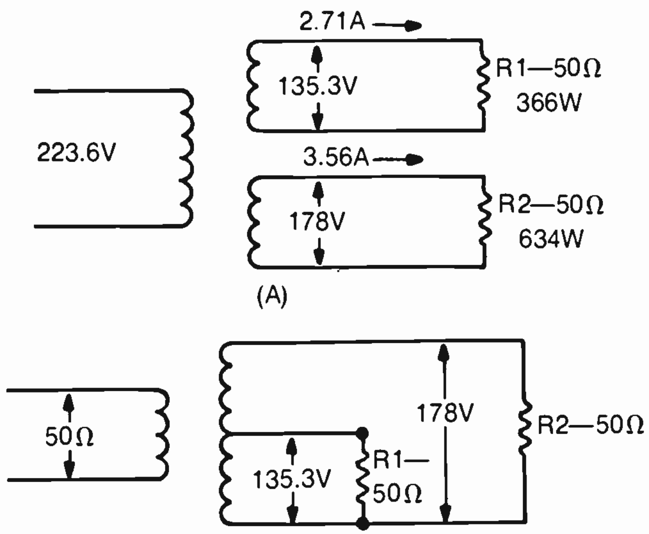


Fig. 11-5. Low-frequency power divider.

RF Power Divider—Series Type

The circuit of Fig. 11-6 is similar to the series power divider shown in Fig. 11-5. In Fig. 11-6 all of the loads again have an impedance of 50 ohms. The problem is to deliver the proper amount of power to each load. Unlike the example of Fig. 11-5, the input to the divider does not necessarily have to have a driving-point impedance of 50 ohms. We can let it assume almost any impedance we wish, then transform this value of impedance back to 50 ohms with an impedance-transforming network. In fact, as we will see, there are good reasons for making the driving-point impedance of the power divider higher than 50 ohms.

The average broadcast engineer isn't interested in designing a power divider. Neither is he interested in just how the component values are arrived at. He is, however, greatly interested in how the various adjustments affect the performance of the antenna array. It is not uncommon to find a power divider with the adjustments so far from their optimum values that it is extremely difficult to get the array into proper adjustment.

The circuit of Fig. 11-6 can be thought of as a parallel tuned circuit with two loads tapped off the coil. As such, it will have a definite value of Q. If the Q is too high the circulating current, and hence the losses, will be high, and the bandwidth will be restricted. If the Q is too low, there will be a lot of interaction between the adjustments.

It is usually easier to *design* a series power divider by starting with the proper position of the bottom tap. The reactance of part of the inductor between the lowest tap and ground should not be less than the characteristic impedance of the transmission line connected to the tap. After a system has been designed and installed, the easiest way to *adjust* it is to start with the setting of the top tap on the coil. This tap should be kept as high as possible on the coil to get the desired amount of current into the line that carries the most power. The positions of the other taps are then set for the proper amount of current in each line. There is a certain amount of interaction

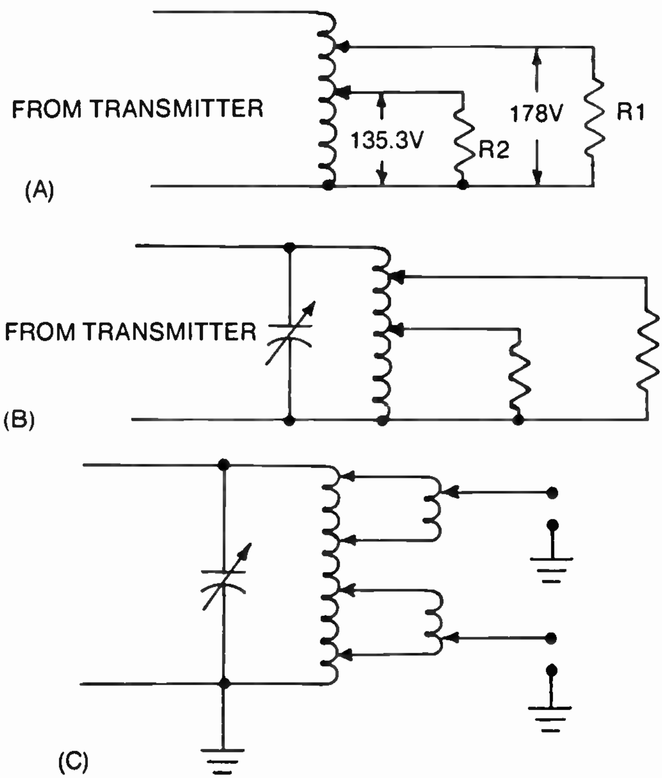


Fig. 11-6. Series power divider.

between the adjustments, and the procedure may need to be repeated several times for optimum adjustment.

As a general rule, the series power divider is well suited for use in arrays where there are more than three towers. The loads are not in parallel, and thus it is possible to keep the driving-point impedance of the divider at a high level. On the other hand, there is apt to be more interaction between adjustments than in the parallel power divider, which we will consider next.

Although a practical power divider may use taps for adjustment (Fig. 11-6B), a more common arrangement is to use *vernier coils* (Fig. 11-6C). Here smaller, continuously adjustable coils are tapped onto the main coil. Since the vernier coils are continuously adjustable, a very precise setting may be made of the effective position of each tap. If an adjustment is pushed to the extreme of its range, it is necessary to move the taps on the coil.

Parallel Power Divider

Going back for a moment to the low-frequency power-division problem that we considered in connection with Fig. 11-5, we could just as well have used a separate autotransformer for each line, as shown in Fig. 11-7. This figure is the basis for the parallel power divider shown in Fig. 11-8. Here a separate coil is used for each transmission line. This arrangement has the advantage of less interaction between adjustments than with the series power divider. On the other hand, since all of the loads are in parallel, it is hard to keep the driving-point impedance high when more than three towers are fed from it.

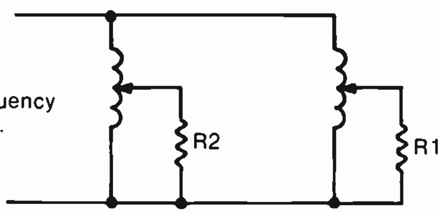


Fig. 11-7. Alternate low-frequency power divider—parallel type.

To summarize, a parallel power divider is easy to adjust but is inefficient when there are more than three towers. In

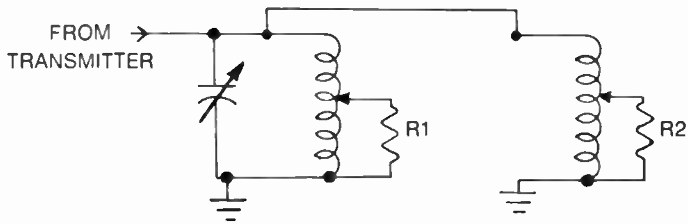


Fig. 11-8. Parallel power divider.

practice, the adjustments are usually made by vernier coils that are tapped onto the main coil, as with the series divider shown in Fig. 11-6C.

To adjust the parallel power divider, the current to the tower that carries the most current is set first. If possible, this tap should be set to the top of the coil, as this will result in the highest driving-point impedance for the divider. For this reason, in some parallel dividers there is no adjustment for the line carrying the highest current; this line is merely connected to the top of one of the coils.

Miscellaneous Power Dividers

Although the series and parallel power dividers described in the preceding pages are by far the most commonly used types, many different arrangements are used, particularly where there are only two towers in the array and the problem of power division is not as complicated. Three such arrangements are shown in Fig. 11-9.

The divider shown in Fig. 11-9A is sometimes called the *unequal-resistance divider*. Here two *L*-networks are used to change the impedance seen at the input of the network and thus the amount of power that the line draws. The power drawn by each line connected through such a network varies inversely with the driving-point impedance of the particular network. Thus if R_1 and R_2 are the driving-point impedances of the two *L*-networks of Fig. 11-9A, the power division will be according to the relationship

$$\frac{P_1}{P_2} = \frac{R_2}{R_1}$$

where P_1 and P_2 are the powers to towers 1 and 2. The input impedance of the two networks in parallel is simply the parallel combination of the two driving-point impedances.

The principal limitation of this type of power divider is that the phase of the current going to each line, as well as the input impedance of the divider, changes whenever an adjustment is made. This means that when we try to adjust the ratio between two currents, we also change their phase, thus necessitating readjustment of the phase control. Although some interaction between ratio and phase adjustments is common, with this particular type of divider the interaction is more pronounced, and proper adjustment of the controls is more difficult. For this reason, this power divider is rarely used on new installations.

Another interesting circuit that has been used for dividing power between two lines is shown in Fig. 11-9B. In this circuit, if the load impedances remain constant, the input impedance will remain constant as long as the capacitor and inductor are varied together. As in the unequal-resistance divider, the power adjustment will also cause a phase shift.

The power divider of Fig. 11-9C takes advantage of the 180° phase difference between the opposite ends of a center tapped

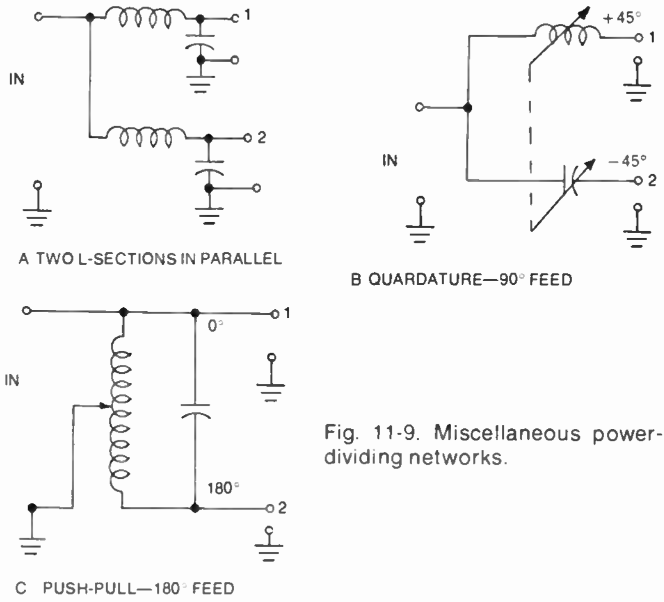


Fig. 11-9. Miscellaneous power-dividing networks.

resonant circuit. It has been used with 2-tower arrays where one tower is much closer to the transmitter than the other, resulting in lines that have widely differing lengths. When the adjustment is near the center of the coil, the magnitude of the currents can be changed with very little phase shift. However, as the adjustment gets closer to one end of the coil, the effect on phase shift will be more pronounced.

COMMON-POINT IMPEDANCE MATCHING

The driving-point impedance of a power-dividing network is usually made as high as practicable. It is thus necessary to have some sort of arrangement to transform this impedance down to a more suitable value for matching to the transmitter. Many different circuit arrangements have been used for this purpose. One common arrangement is shown in Fig. 11-10A. This circuit can be best understood by redrawing it as shown in Fig. 11-10B, with the capacitor broken into two separate units. The input impedance of the coil portion of the power divider almost always has an inductive component, which is tuned out by capacitor C2. The driving-point impedance of this portion of the circuit is then a high resistance. The *L*-network formed by C1 and L1 then transforms this impedance into the desired

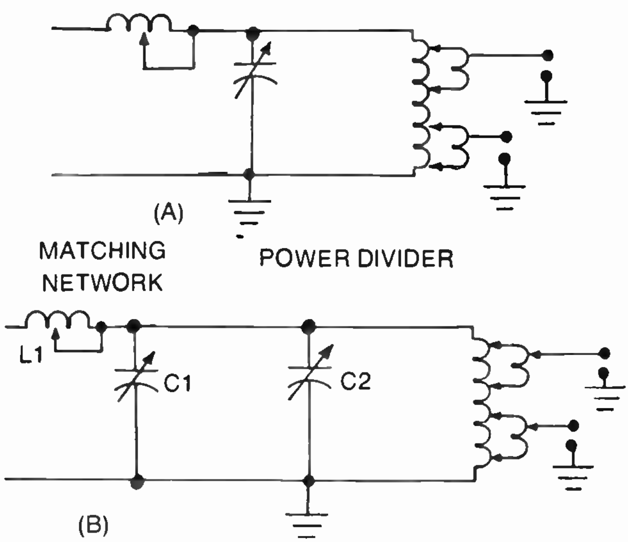


Fig. 11-10. Matching power divider to common point.

common-point value, which is often 50 ohms. As shown in Fig. 11-10A, the two capacitors can be combined and only one physical unit actually installed.

Sometimes the series power divider is connected as shown in Fig. 11-11, with the capacitor at the top of the coil, the input tapped below this, and the various lines connected to taps that are still lower on the coil. The reason for this confusing arrangement becomes clear if the circuit is redrawn as in Fig. 11-11B. Here the two inductors, L1 and L3, can be seen to be the series arms of a *T*-network. Inductance L2 and capacitor C1 form the shunt branch. Thus the circuit is really a *T*-network connected to the tapped coil of a series power divider. Power division is accomplished by the setting of the taps, and the other three taps are used to transform the impedance into the desired common-point impedance.

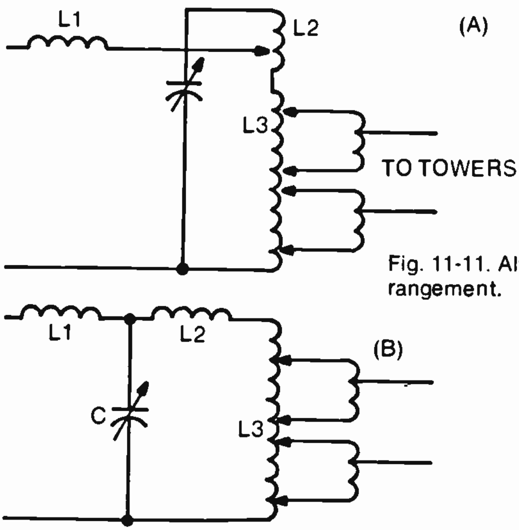


Fig. 11-11. Alternate matching arrangement.

BANDWIDTH

One characteristic of directional-antenna systems that is often neglected is that of bandwidth. The antenna and the feeder systems are designed to operate properly at the carrier frequency. If the design is properly carried out, the system will work equally well out to the highest and lowest sidebands of the signal. This concept is often neglected, however, when an array is adjusted or when design changes are made.

If the bandwidth of the complete antenna system is too narrow, the following effects may be produced:

1. The sidebands may be attenuated or accentuated, resulting in distortion of the signal and reduced radiated power.
2. The phase shift at the sidebands may be substantially different than at the carrier frequency, resulting in a geographical shift of the pattern with modulation. This is particularly noticeable in a pattern with deep nulls.
3. Inasmuch as narrow bandwidth is associated with high-Q circuits, the losses will be high and the efficiency of the system will below.

The most common cause of inadequate bandwidth in an antenna feeder system is the use of a single network to effect an impedance-transformation ratio of greater than 10. A high-impedance transformation ratio results in high-Q circuits, and high-Q circuits inherently have a narrow bandwidth. Narrow bandwidth may also result from the use of short towers, but this is rare in any system installed in recent years, because short towers usually do not provide the minimum field intensity required by the FCC Rules. Sometimes a narrow bandwidth results from the use of a critical array to get a very complex radiation pattern.

In any case, one effect of a narrow bandwidth is the loss of sideband power. There are two factors that contribute to this. First, the impedance seen by the transmitter varies over the bandwidth of the signal so that full modulation cannot be realized at higher audio frequencies. Second, the narrowband system simply will not couple the higher sideband frequencies to the antenna.

When a system is found to have inadequate bandwidth, the cause should be found and, if possible, corrected. Narrowband systems are inherently unstable and will usually continue to cause problems until they are straightened out. One corrective approach that has been taken with limited success is to use the slightly different form of matching network described momentarily.

Figure 11-12A shows coil and a plot of how its inductive reactance varies with frequency. The plot is a straight line; the higher the inductance, the steeper the slope of the line. There are cases where the bandwidth of a network could be improved if we had a reactance whose slope was steeper. Unfortunately the size of the inductor is limited by the design of the network. There is a way that we can get the value of inductive reactance we need and, at the same time, have a reactance that varies faster with changes in frequency than the reactance or a simple inductor does. The scheme is to use a series-resonant circuit in place of one of the inductors in the network. Figure 11-12B shows a series-resonant circuit and how its reactance varies with frequency. Note that both of the circuits in Fig. 11-12 have a reactance of 50 ohms at the design frequency, but that the reactance of the series-resonant circuit changes more rapidly with frequency.

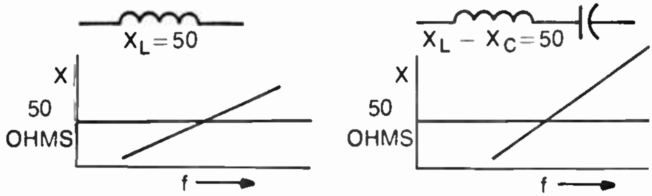


Fig. 11-12. Variation of Reactance with frequency for an inductance and a series LC circuit.

Figure 11-13A shows the usual type of *T*-network used to match a transmission line to a tower. The normalized impedance seen looking into the network is plotted by X_s on a Smith chart in Fig. 11-14. If the network of Fig. 11-3A is replaced by the network of Fig. 11-13B, the result is a normalized impedance as shown by the 0s in Fig. 11-14.

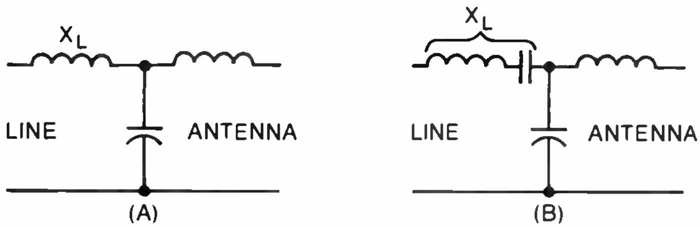


Fig. 11-13. Usual T-network, in A; T-network with series-resonant component, in B.

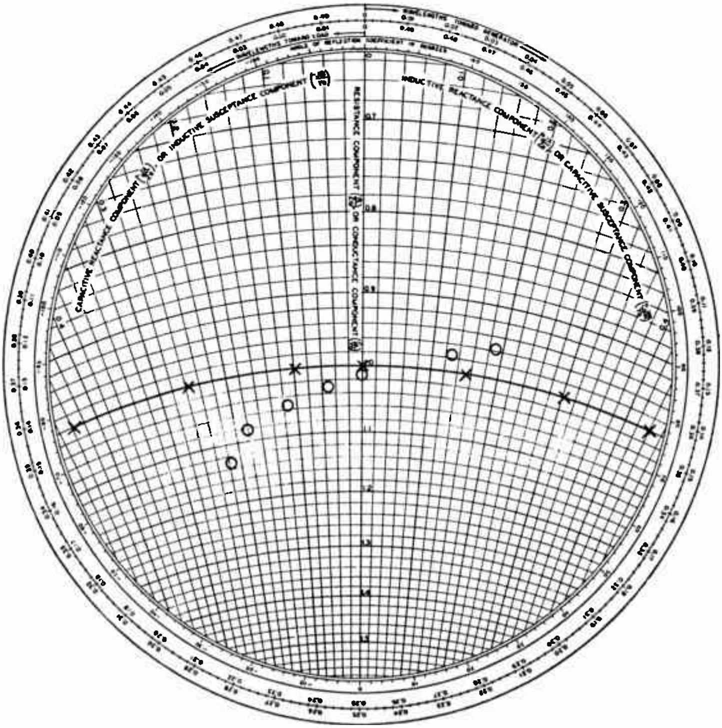


Fig. 11-14. Impedance of networks of Fig. 11-13.

The fact that the points of the new-network plot are clustered close to the prime center of the chart shows that a much more constant load impedance is provided for the transmission line and, hence, for the transmitter. This example is based on an actual occurrence. The towers were too short to be good antennas, but with the modified network, satisfactory performance was obtained.

HANDLING THE NEGATIVE-RESISTANCE TOWER

In an array of four towers or more, the resistive part of the driving-point impedance of one or more of the towers often has a negative value. This means that the tower obtains its energy through the mutual impedance between it and the other towers of the array. This is a confusing situation, but if it is carefully thought out, it will cause no serious problems. We know the following things concerning the negative resistance tower:

1. The tower must carry a current of the proper magnitude and phase.
2. The direction of the current is 180° out of phase with what it would be in a tower having a positive base resistance.
3. We need some method of controlling the magnitude and phase of the tower current.

The simplest, although not the most efficient, way of handling the negative-resistance tower is to terminate it through a matching network to a resistor, as shown in Fig. 11-15. The energy that the negative tower actually gets from the other towers is thus dissipated in the resistor. The magnitude and phase of the current may be controlled by the parameters of the network. Naturally, this isn't a very efficient arrangement, particularly if the negative tower handles a substantial amount of current.

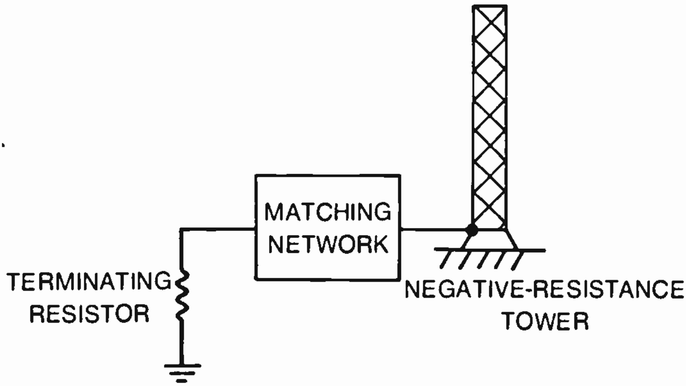
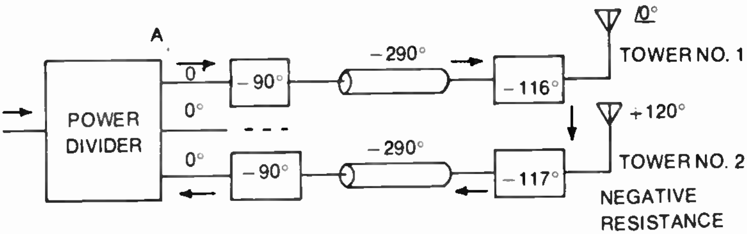


Fig. 11-15. Terminating a negative-resistance tower.

The preferred way to handle a negative-resistance tower is to feed the energy back to the power divider, where it will be passed back into the feeder system again. In this way, all of the energy is radiated rather than some being dissipated in a resistor.

Figure 11-16 shows an arrangement for recovering power from a negative-resistance tower. The procedure is to trace the phase of the signal from the power divider to the first tower, back into the system through the negative-resistance tower, and back to the power divider. When the signal gets back to the power divider, it has the same phase as when it left. That is, it has experienced an integral number of 360° phase shifts. Note that when we add up the phase shifts, we must add or subtract 180° because the current in the negative-resistance tower is flowing in the opposite direction of the current in the other tower.



PHASE SHIFTS FROM A TO NO. 1 TO NO. 2, BACK TO A −90°
−290° −116° + 120−180° −117° −290° −117° = 1080°
−1080° = 360° 3 = 0°

ADD −180° DUE TO CHANGE IN REFERENCE DIRECTION.

Fig. 11-16. Arrangement for handling the negative-resistance tower.

USING ONE TOWER AT TWO FREQUENCIES

Although it would be impracticable with directional-antenna systems, there may be occasions when two transmitters operating at different frequencies in the standard broadcast band use the same tower as an antenna. The problem in this case is to allow the transmitters to feed the tower but not each other. The common way to accomplish this is with resonant circuits.

At *A* in Fig. 11-17 is an *LC* circuit that has two inductors and one capacitor. Branch 1 is inductive at all frequencies. Branch 2 is capacitive at frequencies below its series-resonant frequency and inductive at frequencies above resonance. Thus there will be one frequency below the resonant frequency of 2 where its capacitance resonates with the inductance of branch 1 to form a parallel-resonant circuit, which has infinite impedance (assuming no losses). At a higher frequency (nearer f_h in the figure), branch 2 will be a series-resonant circuit, and its impedance will be zero (assuming no losses). In practice, the impedances will be neither infinite nor zero, but they will be very high at f_l and low at f_h. Thus circuit *A* will shunt any signal at frequency f_h fed back from the tower, but it will not interfere with the progress of signal f_l toward the tower. With the arrangement of circuit *A*, the parallel-resonant frequency will always be lower than the series-resonant frequency.

At *B* in Fig. 11-17 is another circuit with both series and parallel resonance, but with this arrangement, the parallel-resonant frequency will always be higher than the series-resonant frequency (just the opposite of the case of circuit *A*). Consequently, *B* will pass f_L energy to the tower but reject f_h energy. Together, *A* and *B* will pass f_L transmitter, and bypass around the f_L transmitter any f_h energy that is fed back.

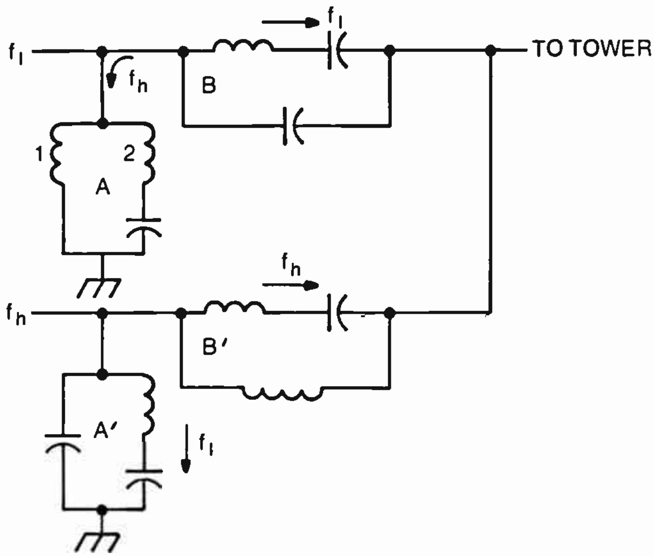


Fig. 11-17. Feeding two transmitters to one tower.

Below, in the figure, circuits *A′* and *B′* will pass f_h energy to the tower and prevent any f_L energy from getting back to transmitter f_h.

By actually following the signal paths in Fig. 11-17, you can see that the energy from both transmitters is fed to the tower, but neither of the transmitters feeds energy to the other.

Chapter 12
Ground System

The theoretical performance of a vertical antenna is derived by assuming that the antenna will be operating over a perfectly conducting earth. This type of analysis is useful in that it will show us the best possible performance that can be obtained from a given antenna. We never get the maximum theoretical performance for a number of reasons, one of which is that the earth is not a perfect conductor.

To compensate to some extent for the effect of the conductivity of the earth on signal propagation, all standard broadcast stations are required to have a ground system. The ground system of a standard broadcast station consists of radial wires extending outward from the base of each tower. Usually these ground wires are buried. It seems that the old adage "Out of sight, out of mind" applies to ground systems. It is common for the ground system to almost never be inspected, and a deteriorated ground system is often responsible for many ills that befall a directional-antenna system, including low efficiency, loss of signal in the primary service area, and general instability of the array.

BASICS OF THE GROUND SYSTEM

Figure 12-1 shows a single-tower broadcast antenna. The electric field from the tower extends from the tip of the tower to the ground. For simplicity the diagram shows only a single line of force. Where the line reaches the ground, a current flows through the ground back to the base of the tower. Thus the lines of the electric field are a part of a closed loop that is completed by the current flowing in the ground back to the antenna.

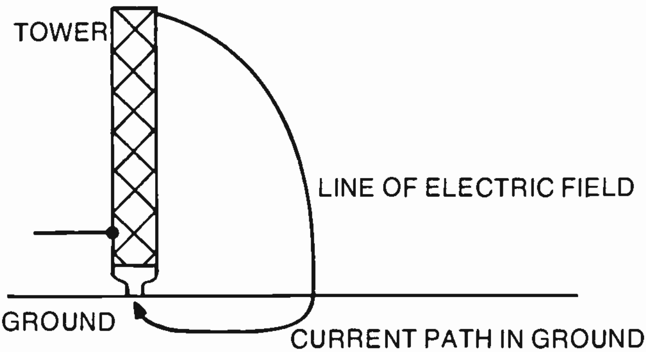


Fig. 12-1. Ground current.

If the earth were a perfect conductor, there would be no resistance to this current, and the lines of the electric field would be perpendicular to the ground at the points where they reached the ground. With an imperfect conductor, which the earth actually is, the situation is different. The lines of the electric field are not perpendicular to the earth, but are actually tilted forward in the direction of propagation (Fig. 12-2). This means that the direction of propagation of the wave is directed slightly toward the earth. This line of propagation can be resolved into two components at right angles to each other. One component, the regular ground wave, is directed along the surface of the earth. The other component is directed downward into the earth and represents the loss that is encountered when a signal is propagated over an imperfect conductor.

Conductivity and Skin Depth

The earth is actually both an imperfect conductor and a dielectric. At broadcast frequencies conduction is the chief phenomenon, and we can usually neglect the dielectric constant. The conductivity of the earth ranges from about 2

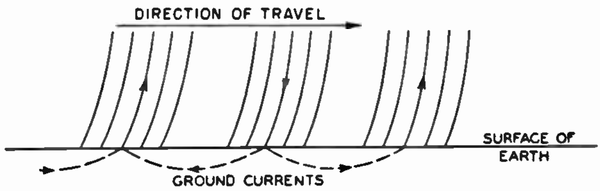


Fig 12-2. Tilting of electric vector of wave front.

mmho/m (2 millimho per meter) for dry, sandy locations to as high as 5 mho/m for sea water. Although the lowest conductivity found on the surface of the earth is about 1 mmho/m, scattering of the wave by rough terrain can introduce losses that will make the conductivity appear even lower than this.

The current flowing in the earth back to the antenna (Fig. 12-2) penetrates the earth for some distance, but the depth of penetration is limited by the skin effect. The *skin depth* is the depth at which the current has fallen to about 37% of its value at the surface. This depth depends on frequency and is smaller at higher frequencies. Nearly 90% of all ground losses occur within this depth. Figure 12-3 shows the variation of skin depth with ground conductivity for various frequencies.

The currents in the ground from all directions come together at the base of the tower. The *ground loss* is the sum of the losses due to all of the ground currents coming from all

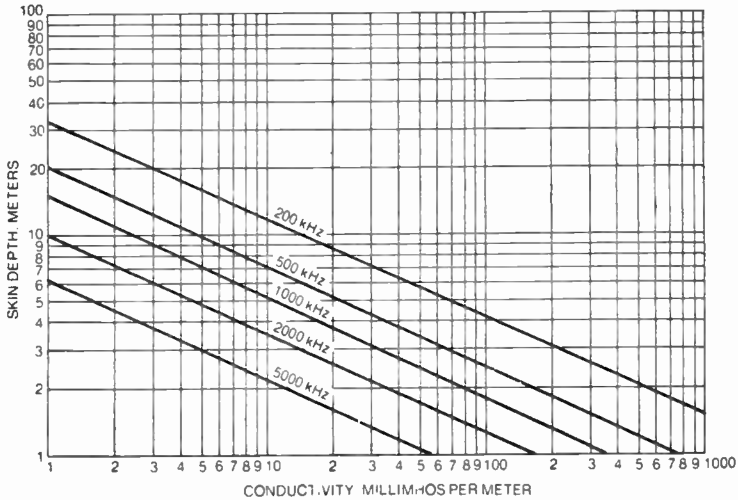
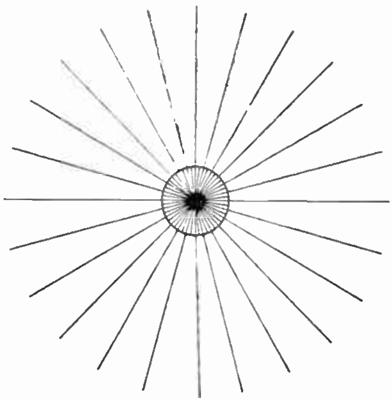


Fig. 12-3. Skin depth versus ground conductivity.

directions. Without some sort of artificial ground system, these losses would be prohibitively high at broadcast frequencies.

Radial Wires

Experimental work has shown that at distances greater than about 1/3 wavelength from the base of the tower, ground losses are almost independent of the tower height. Closer to the antenna the losses increase rapidly as the tower height is decreased. Thus, for antennas of the heights normally used in standard broadcasting, the ground should have a good conductivity out to about 1/3 wavelength from the base of the tower. The usual approach to improving ground conductivity in the vicinity of the antenna is to install a ground system consisting of radial wires extending out from the base of the tower as shown in Fig. 12-4.



240 RADIALS EACH APPROXIMATELY 10 TO 15 DEGREES LONG AND BURIED 2 TO 4 INCHES. 120 RADIALS EXTENDED TO APPROXIMATELY 90 DEGREES AND BURIED 6 TO 8 INCHES ALONG THE EXTENSION.

Fig. 12-4. Ground system for nondirectional antenna.

The current FCC Rules specify that the radials should be at least 1/4 wavelength long and that there should be as many radials as practicable, but in no case less than 90. The Rules add that a system of 120 radials spaced every 3° and extending, 0.35–0.4 wavelength from the tower is considered an excellent ground system. In addition, a square ground screen 24 or 48 on a side is often provided at the base of the tower, particularly when the tower height is such as to cause a high base voltage.

Whenever a less-than-optimum ground system is used, the FCC requires a complete field-intensity survey to establish that the effective field at one mile meets the minimum requirements.

The radial ground system is chosen because the radials follow the natural paths of the ground currents. At one time, ground meshes of crossed wires were thought to form a good ground system. This approach is not used today, because the paths to the base of the tower are not direct, and circulating currents may flow, introducing additional losses.

PRACTICAL GROUND SYSTEMS

The diameter of the wire used for radials doesn't seem to have much influence on the efficiency of the ground system. Systems have been installed with No. 18 enameled wire and, at the other extreme, with 2 inch copper strap. Wire in the No. 10 size is commonly used. In some areas the ground is especially corrosive, and the radial wires deteriorate rapidly. In all cases, the use of ordinary tin—lead solder should be avoided since it usually deteriorates rapidly. Connections should be brazed or made with silver solder.

Some installations have ground rods at the end of each radial, with the outer ends of the radials bonded together as shown in Fig. 12-5. There is some question as to the effectiveness of these schemes except when the radial wires are too short. If the radial wires have the optimum length, the current will be greatest at the tower and will drop to zero at the ends of the radials. The way to tell whether or not bonding or ground rods would be advantageous is to check the current along a radial. If it drops to nearly zero at the end of the radial, there is little to be gained by using ground rods. If, on the other hand, there is still a substantial current at the end of each radial, it may be advantageous to install ground rods and tie the radials together as in Fig. 12-5. The ground rods must be

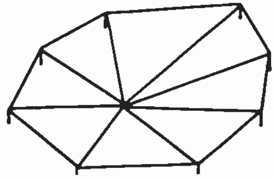


Fig. 12-5. Questionable ground system.

deep to be effective. Referring to Fig. 12-3, we can see that at the upper end of the standard broadcast band, the ground rods would probably have to be over 30 ft long to be effective.

GROUND SYSTEM INSTALLATION

Installing or replacing a ground system is not particularly difficult, but care should be taken to do a good job. The ground system is not something that can be dug up and fixed easily, so it is advisable to do the job properly the first time. Usually a properly installed ground system will do more to make an antenna system stable than any other single factor.

Radial wires of at least No. 10 size should be used, and a ground screen about 24 by 24 ft is recommended. The radials should be plowed into the ground to a depth of about 6 in. and should be as straight as possible. The radials should be brazed to the sides of the ground screen, and any sections in the screen should be brazed together. At least two 2 in. copper straps should run from the bottom of the base insulator to the edge of the screen. The screen should be mounted on a frame at least 4 in. above the ground, and when the installation is completed, the screen should be filled with gravel.

Heavy growths of vegetation in the vicinity of the antenna tower will increase losses. It is generally recommended that the area covered by the ground screen be treated to restrict the growth of vegetation. Used crankcase oil is quite effective for this. The rest of the area over the radials should be seeded, and the area should be mowed regularly to keep the grass low.

When an existing ground system is replaced, there is often a building in the area where it is necessary to run a radial. Tunneling under a building is usually not practicable, so the best practice is to bury a 2 or 4 in. copper strap at a depth of about 4 to 6 in. around the periphery of the building (Fig. 12-6). The radials that would normally lie where the building sits should be run to the strap and brazed to it. The radials can then be continued from the other side of the building as shown.

Finally, in many cases, a ground strap at least 2 in. wide should be run from the ground screen to the transmitter building.

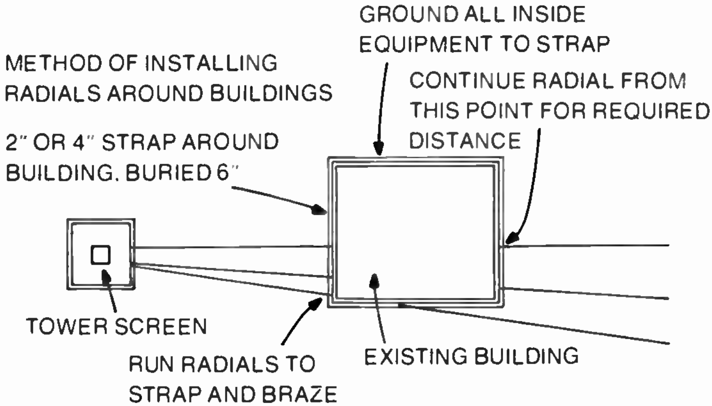


Fig. 12-6. Running ground system past a building.

Wherever possible, all of the radials should be the same length. This isn't always possible, particularly in older installations, where the property belonging to the station may not be large enough to permit equal radials. For example, the antenna may be erected in the center of a rectangular plot as shown in Fig. 12-7A. Then the diagonal radials and those extending along the long dimension of the rectangle may be long enough, but there may not be enough property along the short dimension of the lot to permit radials as long as we would like to have them. The result is often that the pattern of a tower tends to be elongated as shown in Fig. 12-7B. About the only thing that can be done to help this situation is to install long ground rods at the ends of each of the shorter radials. This will probably improve the pattern, but in many cases it still will not be circular.

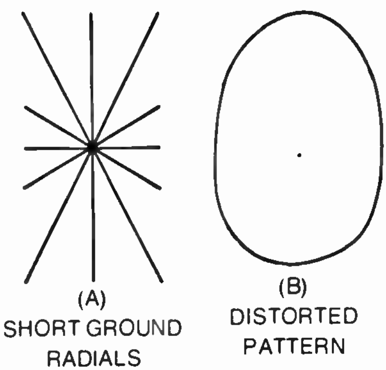
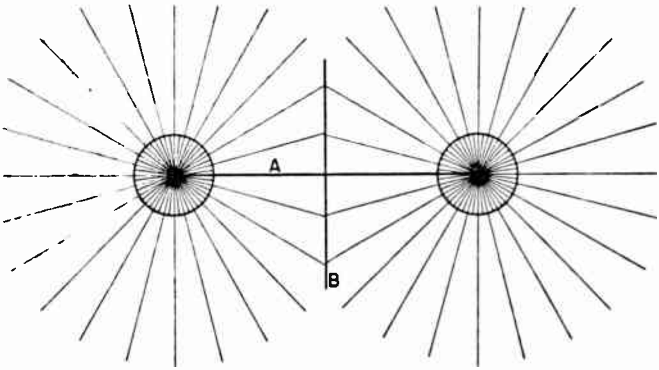


Fig. 12-7. Effect of short ground radials.

DIRECTIONAL ANTENNA GROUNDS

Although the ground system is important with any standard broadcast antenna, it is especially important with a directional-antenna system. The design of such a system is based on all of the towers in the array radiating omnidirectionally with equal efficiency. To accomplish this, each tower of the array needs an effective ground system. With most directional antennas the spacing between the towers is such that the radials from the various towers tend to overlap. This overlapping is undesirable and can be avoided by using the scheme shown in Fig. 12-8. Here 120 radials are installed around each tower. The points where the radials meet are connected together with a copper strap at least 2 in. wide. As with other connections in the ground system, the strap should be brazed to the radials. Another copper strap should be run between the bases of the towers and to the transmitter building.



A — STRAP BETWEEN THE TOWERS
B — STRAP AT POINT OF OVERLAP

Fig. 12-8. Typical directional antenna ground system.

EQUIPMENT GROUNDING

The ground system of the antenna is the place in the station installation that we must consider to be the best ground. Equipment in the transmitter building must be connected to the strap leading from the ground system.

When overhead coaxial transmission lines are used, they should be bonded to the ground system at intervals of not more than 20 ft (Fig. 12-9).

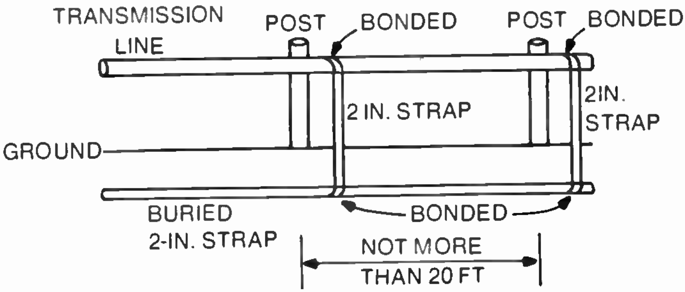


Fig. 12-9. Grounding of overhead transmission lines.

It is almost always necessary to have an impedance-matching network at the base of each tower, and any grounds in this network should be bonded to the ground system. Care should be taken to properly ground any metal enclosures used to house matching networks or base-current meters.

Figure 12-10 shows a satisfactory system of making a ground inside a cabinet. The line is insulated where it enters the cabinet, and a ground connection is made to the inside of the cabinet. In the same way, a strap from the ground system is insulated until it is inside the cabinet, where it is connected to the common ground. With this arrangement all of the currents will flow in the ground conductors on the inside of the cabinet, and there will be no stray currents on the outside surface.

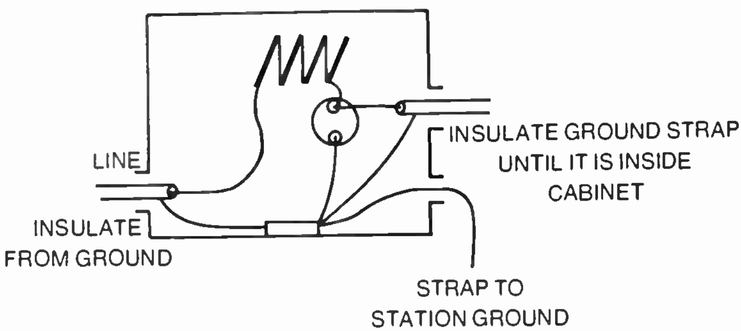
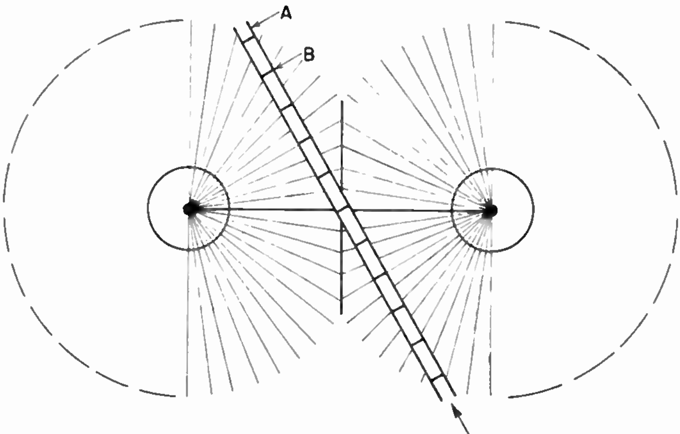


Fig. 12-10. Equipment-cabinet grounding.

SPECIAL GROUNDING ARRANGEMENTS

The ground system is one of the most important parts of the AM antenna system. It is responsible for most of the losses associated with the towers and is a frequent cause of instability in directional systems. The ground system should be the best available, with radials that are long enough to do the job.



A — STRAPS ALONG ROAD OR CREEK BONDED TO RADIAL WIRES
B — BONDS UNDER THE ROAD OR OVER THE CREEK CONNECTING THE STRAPS

Fig. 12-11. Special grounding arrangement to overcome an obstacle.

There will always be situations where obstacles to a good ground system will be encountered. Sometimes, for example, there will be a road or stream running right where a radial should be. In a case like this (Fig. 12-11), the ground system is laid out in the usual way on both sides of the obstacle. Then a copper strap is run along each side of the obstacle. Finally, copper straps are run through the obstacle to complete the circuit. The principles demonstrated in Fig. 12-11 can be modified so as to handle most obstacles in the way of ground radials.

Chapter 13

Antenna Instrumentation and Measurements

There are two classes of instruments used with standard broadcast antennas: (1) instruments used for measuring electrical parameters in the station and on the antenna, and (2) instruments used at some distance from the antenna to determine the intensity of the radiated field. This chapter is concerned with the first of these. Measurements of field intensity and their interpretation are sufficiently different that they are discussed in the next chapter.

The first requirement for any measuring instrument is that its accuracy be better than the tolerance in the device or system it is intended to evaluate. If, for example, an ammeter is used to calibrate another ammeter, the accuracy of the standard should be about ten times better than the accuracy of the meter being calibrated. This principle may seem obvious, but it is frequently violated in broadcast practice. It is not uncommon to find an engineer trying to hold the phase angle of the currents feeding two towers to within 2° with a monitoring system that has an inherent error of 5° or more. Improving the accuracy of measurements in an antenna sytem usually pays off in improved system performance.

There are three quantities commonly measured in an AM antenna system: current, impedance, and, in directional arrays, phase angle.

RF AMMETER

The most common measuring instrument in standard broadcast antennas is the thermocouple RF ammeter. It is used to measure current at the base of each tower, as well as at the common point of a directional-antenna system. Most broadcast engineers feel that the thermocouple RF ammeter is the least reliable instrument in the entire station, and in many instances this feeling is justified. The RF ammeter can, however, be used and read more intelligently if its principles of operation are well understood.

The heart of the thermocouple RF ammeter is a thermal converter, which consists basically of a short heater strip that carries the current being measured, plus a thermocouple to measure the temperature difference between the center of the heater strip and its ends (Fig. 13-1). The operation of the thermal converter is based on these assumptions.

1. All of the current being measured passes uniformly through the heater strip.
2. All of the other parts of the instrument, except the heater strip, are at the same temperature.
3. A constant amount of heat is generated throughout the heater strip by the current being measured.

When the above three conditions are met, the temperature difference between the center of the heater strip and its ends is

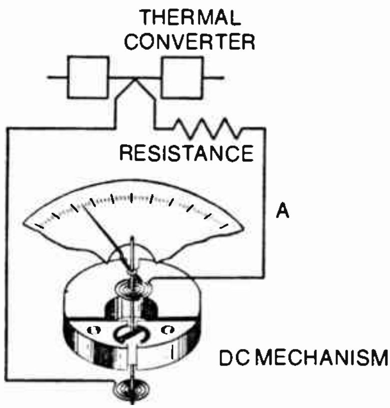


Fig. 13-1. Thermocouple ammeter.

proportional to the *square* of the current being measured. The efficiency of the thermal converter is low, a thermocouple output voltage of 10 mV being typical for full-scale current. This voltage is measured by a conventional D'Arsonval permanent-magnet, moving-coil meter, which is calibrated to read in RF amperes. Whenever full-scale currents of less than about 500 mA are required, the thermal converter is usually housed in a vacuum enclosure to improve its efficiency.

When an RF thermocouple ammeter is used under ideal conditions, it is capable of performing accurate measurements. Unfortunately, the conditions in the average broadcast station are far from ideal. The meter is often located in a "doghouse" at the base of a tower, where it is subjected to temperatures that range from over 100°F in summer to below zero in winter.

One frequent source of errors in RF ammeter measurements is temperature influence. The manufacturer of the meter compensates it so that its indication will be within prescribed limits over a wide range of ambient temperatures. However, the compensation is based on all parts of the meter being at the same temperature, which isn't always the case when the temperature of the meter is changing. For example, when an engineer goes to the dog house to check meter indications, the first thing that he does is frequently something that will change the ambient temperature. In the summer he may open a window or turn on a fan, and in the winter he may turn on a heater. When this is done, the temperature of the meter starts to change. The change of temperature is not uniform; some parts of the meter heat up or cool off faster than other parts. Under this condition the indication may be in error by a large amount. When the temperature of the meter stabilizes—that is, when all parts reach the same temperature—the indication should be within the prescribed limits

One of the enemies of thermocouple ammeters is lightning. It is essential to have an arrangement that will short out the meter when it is not being read. The simple shorting switch of Fig. 13-2A will provide some protection for the meter, but it is not recommended. With this arrangement the meter is not completely removed from the circuit, and the length of the line is not the same when the meter is in the circuit as when it is out of the circuit. In some critical installations this small extra length of line is enough to disturb the indication of the meter.

The arrangement of Fig. 13-2B is preferred. Here the meter is switched completely out of the circuit when it is not in use, and the length of the short is made equal to the length of the circuit through the meter, so that the length of the transmission line will not be disturbed when the meter is switched in or out of the circuit.

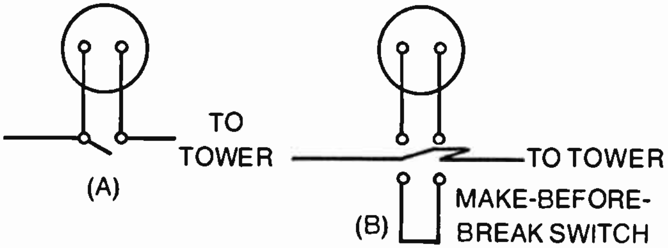


Fig. 13-2. Base ammeter meter protecting switches.

The most likely effect of lightning is that the heater of the thermocouple will be burned out. There is, however, a more subtle problem that can result from current surges due to lightning, even when the meter is switched completely out of the circuit. During a thunderstorm there are many surges of current throughout an antenna feeder system. In a good installation these surges are bled off through air gaps and lightning protectors. Nevertheless, very high currents may flow in the vicinity of the RF ammeter. As with all currents, there are strong magnetic fields associated with these. A transient magnetic field may interact with the field of the magnet in the meter and either strengthen or weaken it, depending on the relative directions of the two fields. If the indication of an RF ammeter changes considerably immediately after a thunderstorm, it is a good idea to check the calibration of the meter before suspecting that the antenna itself is at fault.

Another problem with thermocouple ammeters is that, like other permanent-magnet, moving-coil meters, they are usually calibrated for use on either a magnetic or nonmagnetic panel, but not both. When a meter is mounted on a panel that is made of a magnetic material, some of the flux from the magnet is shunted through the panel, reducing the flux in the air gap of the meter. This causes the meter indication to be low. Manufacturers calibrate a meter for the type of panel on which it is to be used. It is common for a broadcast engineer to calibrate a thermocouple ammeter with it lying on a bench. But if the meter has been calibrated at the factory for use on a magnetic panel, it will be in error when used on the bench.

One more source of error in an RF ammeter is the presence of stray capacitances between various parts of the meter and any RF conductors in its vicinity. As pointed out before, the proper operation of the thermal converter is based on all of the measured current passing uniformly through the heater strip. If a conductor such as a ground wire passes closer to one side of the meter than the other, some of the current in the heater strip may be shunted through the stray capacitance with the nearby wire, resulting in an incorrect indication.

Inasmuch as the RF ammeter is used to measure the operating parameters of a broadcast station, its specifications and application are carefully regulated by the FCC. As of this writing these are two requirements of the FCC Rules that are troublesome:

1. The full-scale indication of the meter must not be greater than three times the normal indication.
2. When not being used, the meter must be stored in a suitable housing at the base of the tower where it is normally used.

These two requirements dictate that different meters must be used to measure the currents in different towers of a directional-antenna array. If one meter could be carried around to each of the towers to measure the base currents, it would probably be easier to keep the ratios between the currents within limits, even if the calibration of the meter were off a little. When separate meters must be used, as is now required, the calibration of each meter must be as accurate as possible to ensure that current ratios will be held within the limits specified in the station license.

Remote RF Ammeter

In addition to the regular base-current ammeters, most stations also use remote-indicating meters back at the transmitter to permit monitoring of base currents. The FCC Rules spell out no less than seven different acceptable ways of remote metering. These are:

1. A second thermocouple may be installed at the base of the antenna, as shown in Fig. 13-3. Here the thermocouple output is fed through an RF filter to keep RF from flowing back along the metering lines. A calibration potentiometer is also included.

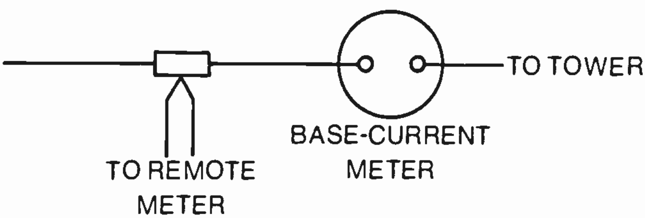


Fig. 13-3. Remote-meter thermocouple must be inserted ahead of base ammeter.

2. The second thermocouple may be inductively coupled to the line (Fig. 13-4).
3. The second thermocouple may be capacitively coupled to the line in a way similar to that used for inductive coupling.
4. A current transformer may be used to couple energy from the line for the remote meter.
5. Remote-control equipment may be used for the remote indication.
6. The antenna-monitoring system may be used for the remote indication.
7. When an antenna is shunt fed and contains only series tuning elements, with no shunt elements, a

transmission-line meter at the transmitter may be used for remote indication of the base current.

In all cases, the coupling to the transmission line at the base of the antenna must be on the transmitter side of the regular base-current meter. Thus any stray currents flowing through the remote-metering system will not flow through the regular base-current meter which, as far as the FCC Rules are concerned, should be the last part before the base of the tower.

The FCC Rules clearly place the responsibility for proper operation of any remote-metering system on the station licensee and the manufacturer of the equipment. Present rules require that the indication of a remote meter be checked against the regular base-current meter at least once a week.

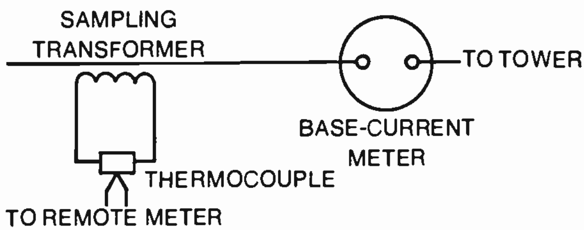


Fig. 13-4. Inductively coupled remote meter.

Calibration of RF Ammeter

Usually the RF ammeters used for standard broadcast work will operate as well on DC or at 60 Hz as at radio frequencies. This characteristic provides a convenient means of calibration. The meter to be calibrated is connected in series with another meter of the same type to an AC source, and the two indications are compared. The problem is getting a meter that is known to be good for calibration purposes. Every station should keep one RF ammeter that can be returned to a calibration house periodically for checking. This meter should be used only as a standard for calibrating other meters. It should be stored where it will not be subject to shock vibration or temperature extremes.

When an RF ammeter is calibrated with direct current, the indication should be checked when the meter terminals are reversed. If there is any difference in indication, it is caused by some of the direct current getting into the millivoltmeter section of the meter.

Modulation Switch

The indication of an RF ammeter in the transmission line to the antenna of an AM station varies with the modulation of the carrier. With 100% sinusoidal modulation the indication will be 22.5% higher than with the carrier alone. Therefore, base currents must be read at instants when there is no modulation on the carrier. It is easier to accomplish this if a switch for removing the modulation from the transmitter is installed at each meter location (see Fig. 13-5). When a current measurement is to be made, the engineer can wait for a pause in the program material, then press the switch for the short time required to make the measurement. In this way, he can be sure that the indication is not being influenced by modulation on the carrier.

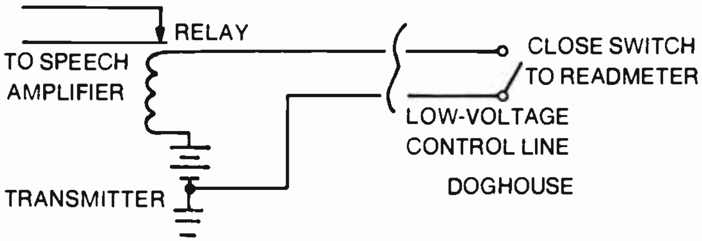


Fig. 13-5. Switch to remove modulation from carrier during antenna-current readings.

RF Current Transformer

One increasingly popular way of sampling RF current on a transmission line or at the base of a tower is to use an RF current transformer. Several of these have been developed over the years. One of the more recent is shown in Fig. 13-6. It consists essentially of a shielded toroidal core that carries the secondary winding of the transformer. The conductor carrying the current being measured passes through the toroid and serves as the primary of the transformer. A current transformer of this type produces a secondary voltage that is proportional to the primary current when the secondary is connected to a fixed value of load resistance. Shielding is provided to minimize coupling through the electric field, so that the coupling between primary and secondary is through the magnetic field only.

The device shown in Fig. 13-6 was designed specifically for use with antenna-monitoring systems. In such applications it is important that not only the transformer ratio but also the phase shift through the transformer be constant. Units of this type typically have a phase-tracking error of about 0.2° from one unit to another.

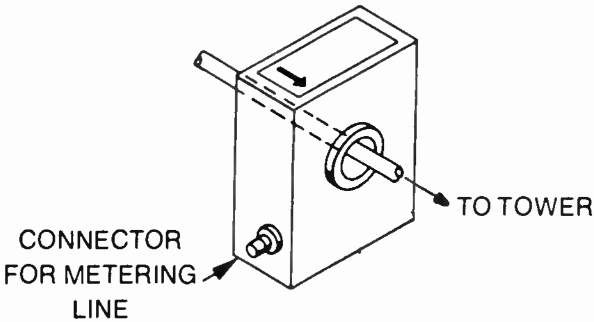


Fig. 13-6. Current transformer for measuring base current. (Courtesy of Delta Electronics.)

The output of the current transformer is connected to an indicator by a circuit such as is shown in Fig. 13-7. Diode D1 in this circuit rectifies the RF signal and provides a DC voltage. A diode of this type has both an offset and a dependence on temperature. In this circuit, diode D3, which is identical to D1, compensates for the effects of changes in temperature in D1. The current required for operating the compensating diode is taken from the signal power by means of diode D2.

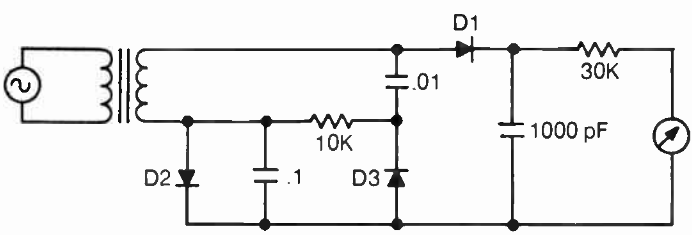


Fig. 13-7. Metering circuit for use with transformer of Fig. 13-6. (Courtesy Delta Electronics.)

The metering circuit of Fig. 13-7 may be used at the base of a tower in place of a thermocouple ammeter, or back at the transmitter site as a remote indicator of base current. In any case, the base-current measurement system must meet the requirements of the FCC Rules and the station license.

ANTENNA-MONITORING SYSTEM

For a directional-antenna array to produce the proper radiation pattern, the magnitude and phase angle of the current in each tower of the array must be held at the proper values. These parameters are monitored by the antenna-monitoring or phase-monitoring system. This system takes samples of the current in each antenna, transmits the samples to a central location, and compares their magnitudes and phases.

Of course, a monitoring system should have an accuracy and stability at least as good as the accuracy and stability of the system being monitored. Unfortunately, in many directional-antenna systems this is not achieved. In fact, there are monitoring systems that are better indicators of wind velocity and ambient temperature than of the operation of an antenna system! The broadcast engineer charged with the responsibility of keeping a directional antenna operating properly must realize that his measurements and adjustments will be no better than his monitoring system. A monitoring system that is not stable should be replaced as soon as feasible. Present FCC Rules and regulations require that the antenna monitor itself be *type approved* by the Commission. At this writing, rules are pending that will tie down the rest of the monitoring sytem as well.

Figure 13-8 shows a block diagram of an antenna-monitoring system. It consists of three parts—a means of sampling the current in each tower, coaxial sampling lines for transmitting the samples to a central location, and an antenna monitor, which compares the magnitudes and phases of the samples.

Sampling Devices

Many different types of devices are commonly used for obtaining samples of currents in the towers of direc-

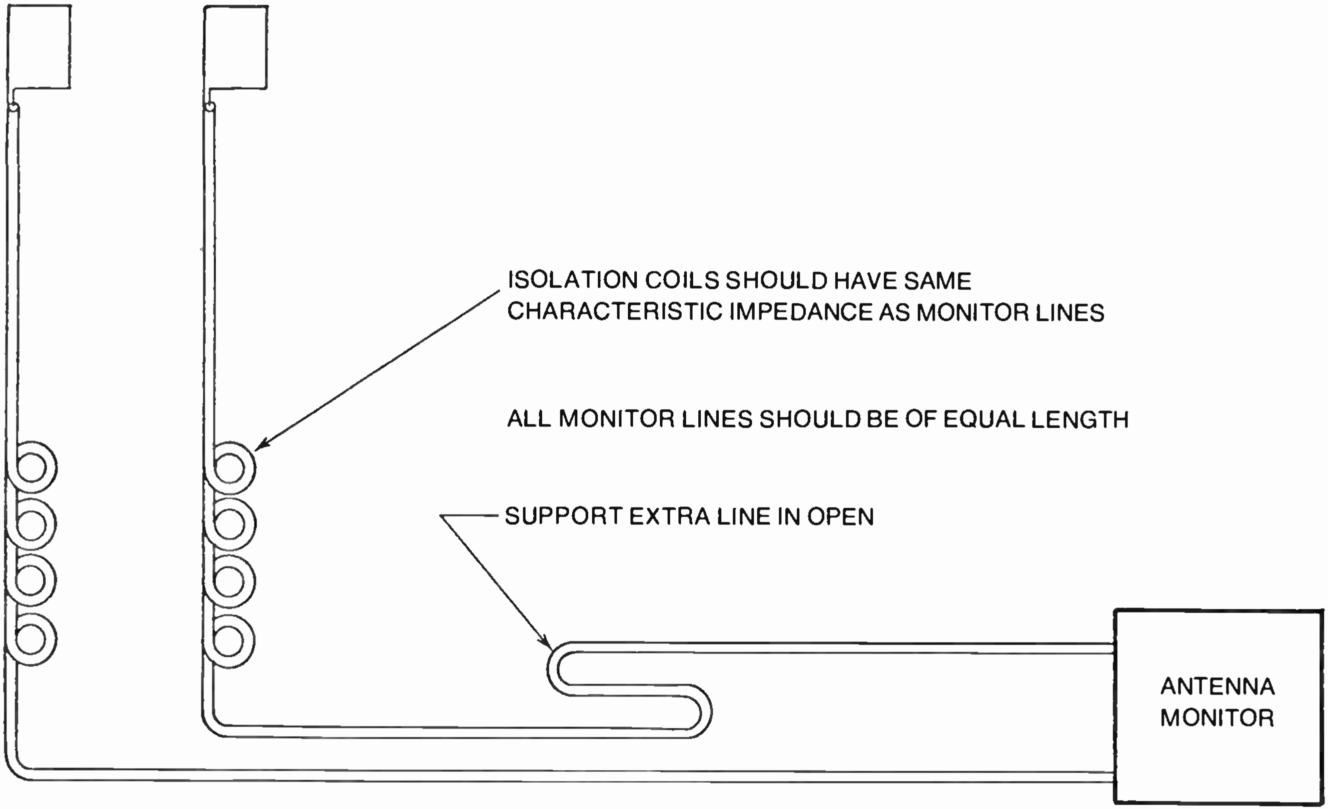
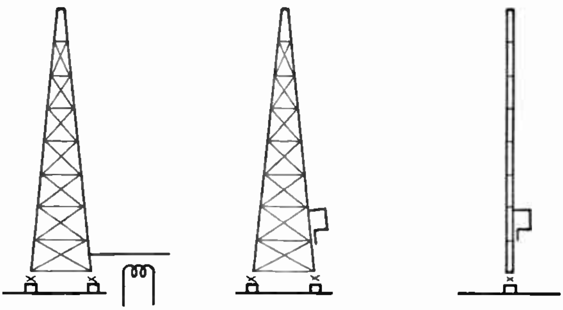


Fig. 13-8. Typical antenna-monitoring system.

tional-antenna systems. Figure 13-9 shows three common arrangements. In Fig. 13-9A a current transformer is installed in the transmission line at the base of each tower. This system works well with arrays that are not critical and where the towers are not much more than 1/4 wavelength in height. This system samples the current in the transmission line, and not in the tower. With a quarter-wave tower the base current is a better indication of the radiated field than with taller towers. With towers that are much more than 1/4 wavelength in height, the relationship between the base current and the maximum current in the tower, and the arrangement of Fig. 13-9A might not be satisfactory.

In the method of Fig. 13-9B a sampling loop is mounted on one leg of a tower, preferably near the place on the tower where the current is maximum. The sample from a loop of this type is much more representative of the radiated field than a sample taken from the transmission line. In some instances, lines for tower lighting and cables for FM or TV antennas mounted on the tower cause the currents in the various legs of the tower to be unequal. The arrangement of Fig. 13-9B tends to sample the current in one leg of the tower. When the currents are unequal, this sample might not be representative of the radiated field. The arrangement of Fig. 13-9C provides a current sample that is more representative of the radiated field than in the other two arrangements.



SAMPLES RELATIONS OF CURRENT AND PHASE IN LEAD-IN
(A)

SAMPLES RELATIONS OF CURRENT AND PHASE IN ONE LEG
(B)

MORE NEARLY SAMPLES CURRENT AND PHASE OF RADIATED FIELDS
(C)

Fig. 13-9. Typical sampling methods.

Of course, broadcast engineers must generally learn to live with an existing monitoring system and seldom have the opportunity to specify exactly what they would like. Nevertheless, the engineer should be aware of the limitations of various sampling arrangements so that, when he has the opportunity, he can modify the system so as to improve its performance.

Many different types of sampling loops are in common use. Some have single turns, others have multiple turns, and some are tuned. In Fig. 13-10A the loop consists of a single turn, with the tower itself forming one leg of the turn. The center conductor of the sampling line is connected to the bottom of the loop, which is insulated from the tower, and the outer conductor is connected to the tower just below the bottom of the loop. Of course, the connections to the tower must be electrically sound and weatherproof.

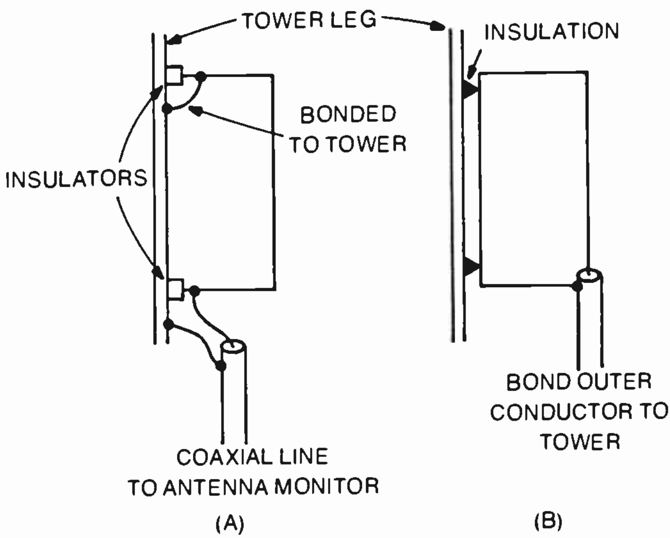


Fig. 13-10. Sampling loops.

The loop shown in B in Fig. 13-10 is similar to that in A except that the loop is complete in itself and is bonded to the tower on one side. In both cases, the loop consists of only one turn and is mounted on a face of corner of a tower so that the plane of the loop passes through the vertical center line of the tower.

Two other factors must be taken into consideration for towers that carry very small currents. (1) When a tower that is a long distance from the central location carries a very small current, there is sometimes not enough current in a single-turn loop to provide a usable sample back at the antenna monitor. (2) A loop on a tower that carries a very small current sometimes picks up more energy from a nearby tower that carries a large current than from the tower on which it is mounted.

The first problem has been handled by using a multiturn loop and actually tuning the loop with a small capacitor to increase its output. With more sensitive modern antenna monitors, this problem is not as severe. A single-turn untuned loop will often suffice. When it doesn't, a multiturn loop is preferable to a tuned loop, which is much more apt to be unstable.

The problem of a loop picking up a signal from another tower is usually solved by proper orientation of the loop on the tower. In the arrangement of Fig. 13-11A, each loop links some flux from each tower. In Fig. 13-11B each loop picks up a signal from the tower on which it is mounted but tends to miss signals from the other towers.

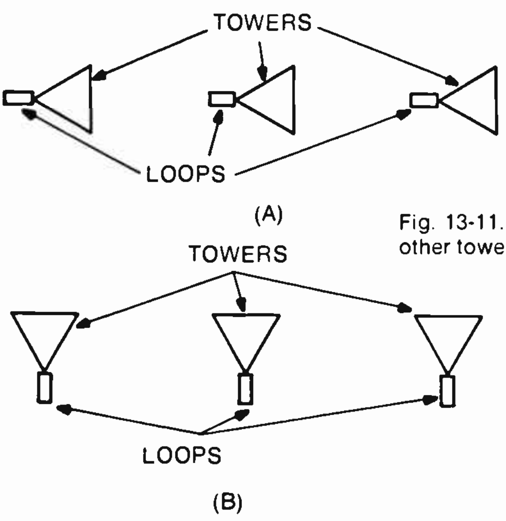


Fig. 13-11. Avoiding pickup from other towers.

Whenever possible, all sampling loops in an array should be identical and mounted in the same way on the tower.

Mechanical stability of the sampling loop is very important. If the position or orientation of the loop changes with the wind, the indications will also change even though nothing in the array itself changes. Heavily galvanized clamps with at least 3/8 in. diameter bolts should be used to secure the loop to the tower, and electrical connections to the sampling line should be made with 1/4 in. brass bolts with nuts and lockwashers. Insulators should be treated to shed moisture. Loops should be at least 10 to 20 above the ground and should not be mounted across a joint in the tower structure.

Sampling Lines

There has been a great deal of debate over the type of coaxial cable to be used for sampling lines. It is generally agreed that the phase stability of RG-type solid-dielectric cables is far inferior to the phase stability of the types of transmission line used to feed the towers. Many engineers have found that a monitoring system behaves much better if the sampling lines are made of polyethelene-foam-filled semirigid coaxial cable. Nevertheless, some systems that are not particularly critical obtain satisfactory results with RG-type sampling lines.

The vertical location of the loop should be the same on all of the towers of the array, and the vertical runs of the sampling cables should be the same length. At the base insulator of the tower, some provision must be made to keep the tower from being shorted to ground through the sampling line. This is usually accomplished by *isolation coils* (Fig. 13-12). These coils may be bought, or they may be made by winding the sampling line itself into a coil. In either case, the characteristic impedance of line through the coils should be the same as that of the sampling line, and all coils in the array should be identical. The outer conductor of the sampling line should be bonded to the tower at least at the bottom of the sampling loop and above the isolation coil. In tall towers, where the vertical run of sampling line is long, it should be bonded to the tower more frequently.

The sampling lines from each tower to the central location should all be the same length. If the distances from the towers

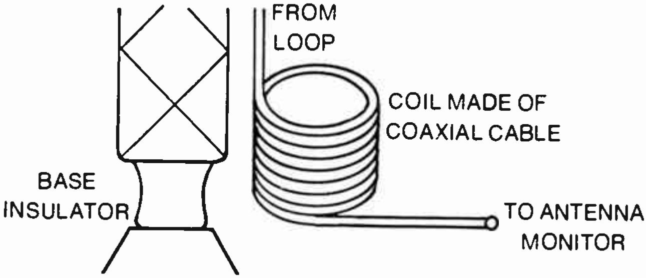


Fig. 13-12. Isolation coil for sampling line.

are not equal, which is usual, the excess line should be coiled up where it will be exposed to the same temperature as the other lines. Excess line from the closer towers should not be coiled up inside a building, where it would be subjected to a different temperature than the other lines experience.

Checking the Sampling System

There are two important considerations in any sampling system. First, the system should be as stable as possible. When a change in the ratio of two currents or in their relative phase is noted, it is nice to have a reasonable assurance that the change has taken place in the array itself, and not in the sampling system. Second, enough data should be available for the engineer to determine through tests whether his sampling system has changed and to get it back to its original condition.

On each portion of a sampling line the following measurements should be made and recorded:

1. The DC resistance of the line should be measured with the far end alternately open and shorted. In the open condition the resistance should be very high, and in the shorted condition it should be very low.
2. The RF impedance of the complete line with the loop attached should be measured at some frequency where the impedance is high. This measurement will help to spot changes that have occurred in the line and loop over a long period of time.

The subject of measurements of sampling systems is under consideration by the FCC at the present time, and more meaningful measurements will be developed and eventually

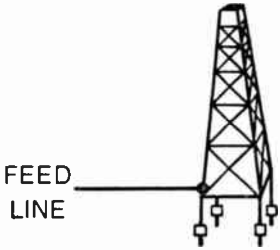


Fig. 13-13. More current will flow in one tower leg than in the others in this tower.

spelled out in the FCC rules. In general, any measurement that can help the engineer detect changes that might have occurred in his sampling system will be well worth making.

Tower Current

The purpose of the base ammeter is to measure the current flowing into the base of the tower. The purpose of the sampling loop is to measure the current at the current loop of the tower. Ideally, there is a good correlation between these two currents but for this to be true, the current must be uniform throughout the tower. This means that the connection to the tower must couple the current equally to all of the legs. Usually this isn't a problem with uniform towers, but towers such as the one shown in Fig. 13-13 that have four legs and four base insulators may present problems. If the feed line is connected to one of the legs of the tower in Fig. 13-13, the current tends to be unequal in the four legs. The best way to feed a tower of this type is to bond copper straps to the legs (Fig. 13-14). The straps are bonded together at the point where they cross, and the feed line is connected at this point.

Antenna Monitor

The purpose of the antenna monitor, which was formerly called the *phase* monitor, is to accept the samples from the

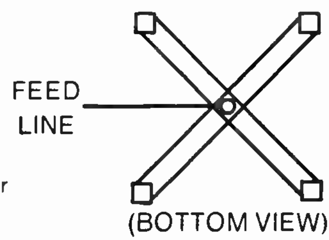


Fig. 13-14. Feed arrangement for 4-leg tower.

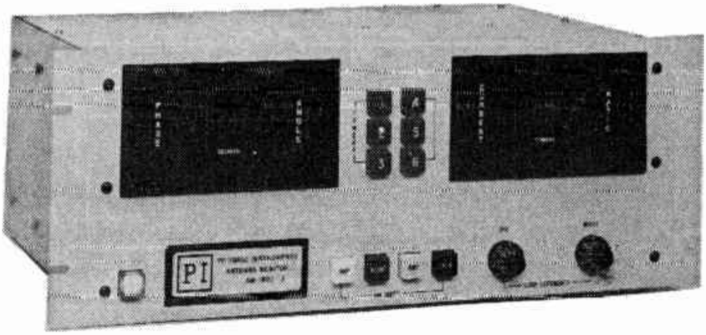


Fig. 13-15. Potomac Instruments' model AM-19D(210) digital antenna monitor.

sampling system and provide indications of the ratios and phase angles between tower currents. The engineer compares the readings with the corresponding values specified in the station license to assure that the operating parameters of the station are within prescribed limits. Current FCC Rules require that the antenna monitor be type approved by the FCC.

Antenna monitors with either analog or digital readouts are available. Figure 13-15 shows a modern antenna monitor with a large digital readout of both ratio and phase angle. Sampling currents from up to 12 towers can be handled by such an instrument. A particular tower is selected by means of a pushbutton control, and the instrument displays the current ratio and phase angle with respect to the reference tower of

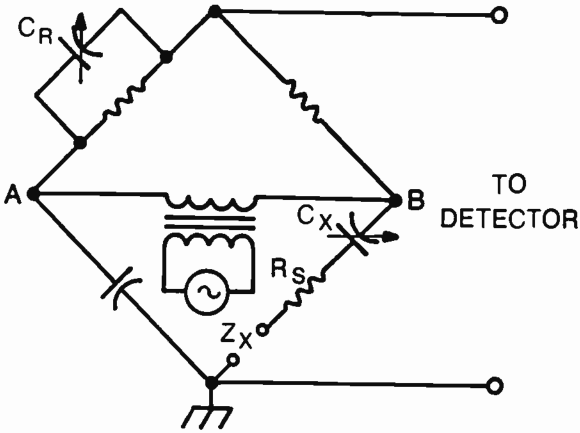


Fig. 13-16. Impedance bridge.

the array. Modern antenna monitors are compatible with remote-control equipment so that a display of current ratio and phase can be obtained at the remote-control location.

IMPEDANCE MEASUREMENTS

The values of resistance and reactance of towers of antenna systems are important parameters, as are the values of the components used in matching and phase-shifting networks. The broadcast engineer must have a means of accurately measuring these values. The common instrument used for measuring impedance at radio frequencies is the radio-frequency impedance bridge. A typical circuit is shown in Fig. 13-16. A signal source such as a calibrated signal generator is connected to terminals *A* and *B* in the figure. The impedance being measured is connected across the terminals marked Z_X . A null detector is connected across the bridge.

The bridge is balanced for both resistance and reactance measurements by variable capacitors rather than variable resistors. The precision variable capacitor is superior to most variable resistors in that is has better linearity, better accuracy, and lower contact noise. To measure resistance, capacitor C_R is varied until the output of the null detector is minimum. The dial of C_R is calibrated directly in ohms of the resistance component of the unknown impedance. Reactance is measured by varying C_X until the bridge is nulled. For reactance, the dial of C_X is calibrated in ohms at 1 MHz. At other frequencies a correction factor must be applied to the reading of the reactance dial.

Inasmuch as the operation of the impedance bridge depends on detecting a null, all components must be well shielded, and any leakage must be balanced out. Balancing out the leakage is difficult with a bridge circuit because the circuit is unbalanced with respect to ground. Stray capacitances are balanced out initially by a balance control, which must be properly adjusted at each frequency at which measurements are to be made.

Usually a well shielded signal generator is used for a signal source, and a communications-type receiver that will tune the standard broadcast band is used for a null detector.

Connections between the bridge and signal source are made with coaxial cable, and all components are well grounded. The signal source must be calibrated accurately because the accuracy of the measurements depends on an accurate knowledge of the frequency at which the measurements are made.

Measurements on a distributed element, such as an antenna tower or transmission line, are complicated by the element itself acting as a receiving antenna. If a signal from a cochannel station is picked up, it will make the null very broad and hence will decrease the accuracy of the measurement. The interfering signal does not have to be very strong to seriously affect the accuracy of an impedance measurement. At the point of measurement, the signal from the source is intentionally set at a minimum, so a weak interfering signal will have a noticeable effect. The use of a receiver as a null detector makes it easy to tell when a interfering signal is strong enough to influence the accuracy of a measurement. When the bridge is balanced, the signal from the signal source will be at its minimum, and any interfering signal will be heard in the speaker of the receiver.

Fig. 13-17 shows a scheme that can be used to measure the impedance of a load under operating conditions. Here a series resonant circuit, consisting of a capacitor and a tapped coil, is connected between the antenna ammeter and the load. If possible, the tap position on the coil should be continuously variable. Naturally, both the coil and the capacitor must be capable of carrying the full operating current.

If the load has a capacitive component, the coil is placed closest to the antenna as shown in Fig. 13-17A. The positions of the coil and capacitor are reversed if the load has an inductive component, as shown in Fig. 13-17B.

The voltmeter is a high impedance RF voltmeter. Its impedance must be very high compared to the value of the load impedance.

The first step is to connect the voltmeter directly across the load. The load voltage (V), together with the load current (I), will permit us to compute the magnitude, but not the phase angle, of the load impedance from the relationship

$$|Z| = \frac{V}{I}$$

The next step is to slide the voltmeter along the coil until a point is found where the voltage indicated on the meter is minimum. At this tap position, the reactance to the right of the tap is equal and opposite to the reactive component of the load. The circuit to the right of the tap, including the load impedance, forms a series-resonant circuit. Thus the impedance seen looking to the right from the tap position is equal to only the resistive component of the load impedance; the reactances cancel out. This makes it possible for us to compute the resistive component of the load from the equation

$$R = \frac{V_{min}}{I}$$

where V_{min} is the voltage at the position of the tap where the indication of the voltmeter is minimum, and I is the series current. If a high impedance voltmeter is used, the value of the series current I should be the same for both measurements.

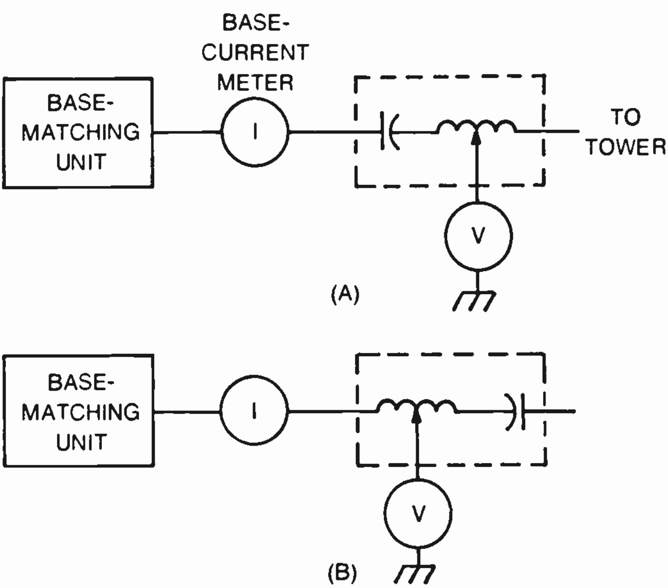


Fig. 13-17. Method of measuring driving-point impedance.

We now know the magnitude Z of the unknown impedance and the value R of its resistive component. We can compute the magnitude of the reactive component X as well as the phase angle γ from the equations

$$X = \sqrt{Z^2 - R^2}$$

$$\tan \gamma = R/X$$

Operating Impedance Bridge

The *operating impedance bridge*, or *OIB*, permits measuring the complex impedance across a load under operating conditions. In directional-antenna systems, where the driving-point impedance of a tower is considerably different from its self-impedance, this instrument is invaluable. The instrument can operate with considerable power (up to 5 kW) being supplied to the load. This high power level permits the use of an internal null detector rather than a sensitive communications receiver and thus avoids many of the problems encountered with low signal levels. All of the components of the operating bridge are in parallel to ground, so that problems of stray capacitance encountered with a conventional impedance bridge are avoided.

Figure 13-18 shows a simplified schematic of the operating impedance bridge. The circuit between the generator and load is interrupted by a short (typically about 9 in.) section of transmission line with a characteristic impedance of about 150 ohms. Because this section is so short, it has practically no effect on most antenna impedances found in standard broadcast stations. A second section of transmission line is

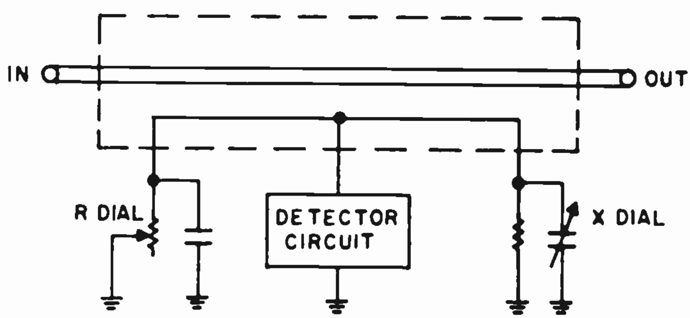


Fig. 13-18. Simplified schematic of operating impedance bridge. (Courtesy Delta Electronics.)

lightly coupled to the main section. This second line has waves proportional to the forward and reverse waves in the main section, which in turn are a function of the resistive and reactive components of the load impedance. The measurement is accomplished by adjusting resistance and reactance dials for a null indication. The resistance dial is calibrated directly in ohms of resistance. The reactance dial is calibrated in ohms of reactance at 1 MHz. For measurements at other frequencies, a correction factor is applied to the dial indication.

Problems of cochannel-station interference are minimized with the operating bridge by using sufficiently high power levels to minimize the effect of weak interfering signals.

The operation of the operating bridge is straightforward. It can, of course, be used with a signal generator and a null detector, just like a conventional bridge. Then, however, leakage is just as important as in any other bridge measurement.

When we are making bridge measurements at high power levels, we must exercise special caution. A short circuit at 5 kW is usually much more expensive and hazardous than a short circuit at the signal levels of a signal generator. The case of the operating bridge must be properly grounded; otherwise, very high voltages may be encountered.

Although the operating bridge can conveniently perform many impedance measurements that are impractical if not impossible with conventional bridges, a conventional precision impedance bridge is still needed in a standard broadcast station for precise measurement of the values of the components in the feeder system.

A rather problematical impedance in a directional-antenna system is the impedance at the *common point*. The power radiated by the station depends on the value of the common-point impedance, and this value can change with adjustments in the feeder system. A special version of the operating impedance bridge, called the *common-point bridge*, is available for permanent installation at the common point. Bridges of this type can handle up to 50 kW of power.

Provision is made in the common-point bridge for installing an RF ammeter. This permits measuring the common-point impedance at any time—even when the station is operating at full power—and continuously monitoring the common-point current. This in turn permits operating the station at its full authorized power at all times. Furthermore, periodic measurement of the common-point impedance often permits detecting changes in the other operating parameters of a directional antenna.

Base-Impedance Measurement

The driving-point impedance seen at the base of a tower is an important parameter in any AM antenna system. It must be known accurately and measured whenever there is reason to suspect that any of the parameters of the antenna system have changed.

In nondirectional antennas, where there is only one tower, the driving-point impedance at the base of the tower is the same as its self-impedance, and the measurement is straightforward. In multitower directional antennas the two impedances are not at all the same thing, and the measurement is much more involved.

An impedance is the complex ratio of the voltage across two terminals to the current flowing in them. Thus the impedance seen at the base of a tower depends on the current distribution on the tower itself. Anything that might affect the current distribution can change the base impedance. Leads from tower-lighting circuits and sampling loops are the items most likely to disturb the current distribution and hence the base impedance. Many engineers feel that it is advisable to feed tower-lighting circuits through an Austin transformer, particularly with directional antennas, because transformers seem less likely to disturb the base impedance.

Self-Impedance

The *self-impedance* of a tower is the impedance that would be measured across the base insulator with no other towers in the vicinity. This is, in fact, what is measured by connecting a bridge across the base insulator of a nondirectional antenna.

In a directional antenna the driving-point impedance seen across the base insulator of a tower depends not only on its self-impedance but also on the mutual impedance with other towers and the ratios of the currents in the towers. Considering only two towers. the driving-point impedance is given by the equation

$$Z_1 = Z_{11} + \frac{I_2}{I_1} Z_{12}$$

where Z_1 = driving-point impedance at base of tower 1
Z_{11} = self-impedance of tower 1
Z_{12} = mutual impedance between towers 1 and 2
I_1 = current in tower 1
I_2 = current in tower 2

Inspection of this equation shows that there was some way to make the current I_2 in tower 2 zero, the driving-point impedance of tower 1 would be equal to the self-impedance; that is, the second term in the equation would equal zero. This can, in fact, be done, although it isn't always easy. The trick is to connect the base of tower 2 in such a way that little if any current will be induced in it from tower 1. If tower 2 is a quarter-wave tower, we can make its current minimum by merely floating it above ground. If it is a half-wave tower, we can reduce its current by grounding its base. If the length is some other fraction of a wavelength, we can minimize the current by connecting a series- or parallel-resonant circuit between the base of the tower and ground and tuning for minimum current. This adjustment is made by driving one tower with low power while adjusting the tuning networks for minimum current in the other towers (Fig. 13-19).

The driving-point impedance of any tower in a directional-antenna array can be measured directly by the method of Fig. 13-17 or by using an operating impedance bridge. As pointed out earlier, the operating impedance bridge has the advantage of permitting base-impedance measurements with the tower operating at substantial power levels.

Impedance measurements are very important in standard broadcast stations. When power is determined by the direct

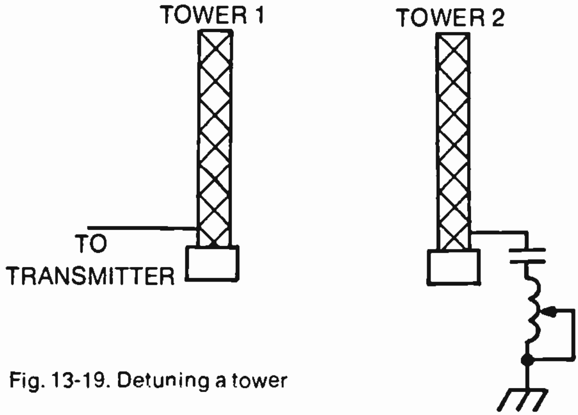


Fig. 13-19. Detuning a tower

method, the reading of the base-current meter is squared and multiplied by the base resistance to determine the operating power. In a directional-antenna system the same procedure is used, but both the impedance and current are measured at the common point. Base and common-point impedance measurements should be made not only at the carrier frequency but also at several frequencies on either side of it.

Figure 13-20 shows a plot of resistance and reactance measurements made at the common point of a directional-antenna system. This plot varies widely from one

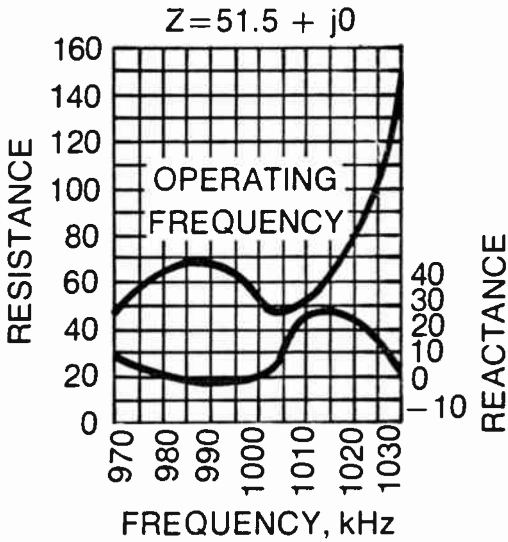


Fig. 13-20. Plot of common-point resistance and reactance.

directional antenna to another and is an indication of the bandwidth of the array.

Measuring Mutual Impedance

The average broadcast engineer generally has no occasion for measurement of the mutual impedances of an antenna array. About the only time such measurement is called for is when an array is first installed or when major modifications are made. The measurement requires an ordinary impedance bridge or an operating bridge. The procedure will be explained next, using a 2-tower array to keep things from becoming too complicated. It can readily be extended to the multitower case.

The first step is to measure the self-impedances at the bases of the two towers with the current in the other towers brought to zero by either opening them at the base or detuning them. From these measurements we get the resistive and reactive components of the self-impedances of both towers.

$$Z_{11} = R_{11} + jX_{11}$$
$$Z_{22} = R_{22} + jX_{22}$$

We need one more quantity to compute the mutual impedance. We get it by measuring the impedance of tower 1 when tower 2 is series-resonated. Tower 2 is series-resonated by connecting the base of tower 2 to ground through a reactance that is equal and opposite to the reactive component of its self-impedance (Fig. 13-21). We will call the impedance that we measure at tower 1 under this condition Z_{11}'. We can now compute the mutual impedance from the equation

$$Z_{12} = \sqrt{R_{22}\,(Z_{11} - Z_{11}')}$$

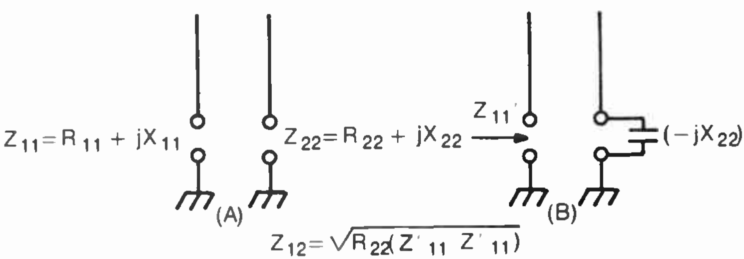


Fig. 13-21. Mutual-impedance measurement.

To solve the above equation, we must extract the square root of a vector quantity. This is done by taking the square root of the magnitude and dividing the angle by 2. Thus the square root of $4/30°$ is $\pm 2 \angle 15°$. The $\pm$ means that there are actually two separate square roots. We can write $-2 \angle 15°$ as $+2 \angle 15° + 180°$ or $2 \angle 165°$. Thus the two values of the square root of $4 \angle 30°$ are $2 \angle 15°$ and $2 \angle 165°$.

When we determine the mutual impedance between two antennas using the method just described, we obtain two values. To find out which one is correct, we can use the mutual-impedance graph in Chapter 9. The values in the graph will be close enough for us to distinguish which of the two values is correct.

Chapter 14

AM Field Intensity Measurements

Probably the most important measurement that a broadcast engineer is called on to make is the measurement of the field intensity of the signal from a directional antenna. Measurements must be made before a station license can be issued, as well as at regular intervals to assure that the pattern of an antenna hasn't changed from the values specified in the license. Not only is the measurement itself important, but proper interpretation of the measurement is equally important.

FIELD-INTENSITY METER

Fig. 14-1 shows a typical field-intensity meter of the type used at standard broadcast frequencies. It is essentially a calibrated receiver with a loop antenna and a meter to indicate field intensity.

There are many reasons for using a loop antenna for field-intensity measurements. For one, it is possible to calculate the intensity of the field directly from the voltage induced in the loop antenna. This avoids the problem of determining the effective height, which we would have to do with some other types of antenna, such as a rod. Another advantage of the loop is its directionality. This permits us to be sure that the field being measured is received directly from

Fig. 14-1. Field-strength meter FIM-41.

the antenna being investigated and is not reflected from some other object.

Figure 14-2 shows a sketch of a loop antenna. Note that the direction of maximum pickup is along the plane of the loop and that there is a null in the direction perpendicular to the plane of the loop. In the instrument of Fig. 14-1 the loop antenna is mounted in the lid of the case and is connected to the input of the receiver when the lid is opened. The instrument has a calibrated attenuator, which permits measurements to be made over a wide dynamic range. Usually a calibrated oscillator is included to permit periodic checking of the calibration.

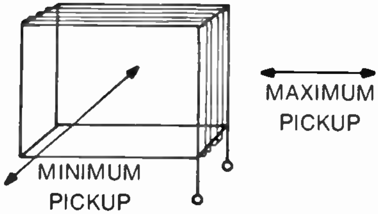


Fig. 14-2. Basic loop antenna.

Some field-intensity meters are crystal-controlled at the carrier frequency of the station to avoid the requirement for accurate tuning. Others have a tuning range that includes the second harmonics of the standard broadcast band to allow measurement of the ratio of the second harmonic to the carrier strength. Highly selective filters are included to minimize errors from stations operating on adjacent channels.

SELECTING THE MEASUREMENT SITE

The first step in obtaining an accurate indication of the radiation from an antenna is to find a suitable location for the measurement. Conducting objects tend to disturb the field from an antenna and must be avoided when field-intensity measurements are made. Nearby conducting objects such as power lines, poles, signs, and similar structures within several feet of the measurement point can cause errors. As a rule of thumb, the distance from a conducting object to the spot where the measurement is made should be at least five times the height of the object, and preferably ten times.

Other potential sources of field disturbance at a proposed measuremnt site are water and gas pipes, power lines, and other underground conductors.

There are two other requirements for a site for accurate field-intensity measurements. For one, the site must be accessible in all kinds of weather so that the measurement may be repeated at any time. A site that isn't accessible because of deep snow in the winter or flooding in the spring is of little value. To check on whether or not an antenna pattern has changed, it is necessary to make measurements at exactly the same place where the original measurements were made.

Additionally, the measurement site must be clearly identifiable. A clear and easy-to-follow description must be provided. If practicable, a photograph should be taken so that the measurement site can be easily found and identified.

The quality of a potential measurement site may be checked with the field-intensity meter. First of all, the direction of arrival of the signal, as indicated by the orientation of the loop, must be in the actual direction of the station. If the signal appears to arrive from a different direction, the proposed site is probably no good. Another very useful check on the quality of a site is to turn the loop antenna of the meter so that the null is in the direction of the station. The null should be sharp and deep. If the null is broad, the site is probably not suitable for accurate measurements.

Another check is to make several measurements within a radius of 5 to 10 ft at the measurement site. The measured field intensity should be essentially constant all around the measurement site. If there are any shape changes, select another site. Even if one point in the area provided an accurate measurement, it would be nearly impossible to find the same spot for a subsequent measurement.

Once a measurement site is selected, it should not only be identified so that it can be found easily in the future, it should be described well enough that newly constructed buildings, power lines, and the like that might render the site useless can be readily identified. It is a good scheme, whenever a measurement site is selected, to also select an alternate site. Then, if for some reason the original site becomes useless, the alternate site is available for future measurements.

INVERSE FIELD AT ONE MILE

The basis for all standard broadcast allocations is the effective or inverse field at one mile from the antenna. It is important to note that this is *not* the field intensity that we would measure if we were to take a field-intensity meter out one mile from the antenna and make a measurement. Before considering just how we go about finding the inverse field at one mile, let us review how field intensity varies with distance from the antenna.

If the earth were a perfect conductor, the field intensity would vary inversely with the distance from the antenna.

Salt water is a nearly perfect conductor, for all practical purposes, but the earth itself is far from it. Suppose, for example, we found that the field intensity from a station was 200 mV/m at a distance of 10 miles from the station over sea water. If we made a similar measurement 20 miles from the antenna, still over sea water, we would find a field intensity of 100 mV/m. In other words, if we double the distance from the station, we find half the original field intensity. Now, over the ground the field intensity attenuates much more rapidly than over sea water because the earth is not a perfect conductor. Much of the signal is dissipated in the resistance of the earth. The amount of attenuation depends on the conductivity of the earth and the frequency of the signal. For the same value of conductivity, the attenuation will be greater for higher frequency signals.

The conductivity of the earth varies from as little as one millimho per meter in dry, sandy locations to as much as 10 or 20 mmho/m in areas where the soil is moist. The ground conductivities at various parts of the continental U.S. are given in Fig. 6-10.

The inverse or effective field at one mile from the transmitting antenna that is used as the basis of allocation is the field that would exist *if the ground were a perfect conductor*. By making a series of measurements along a radial from the antenna and plotting them properly, we can find not only the effective field at one mile but also the effective value of ground conductivity in the vicinity of the station.

RUNNING THE RADIALS

To determine the effective field at one mile from the antenna, we must start with a series of measurements taken along radial lines extending out from the reference point of the antenna array. It is essential to have an accurate map of the area. The maps issued by the U.S. Coast and Geodetic Survey are the best for the purpose. If these are not available, state or county maps may be used, but the distances should be checked with an automobile odometer. Road maps often have large errors and should not be used unless they have been completely checked.

The first step in making radial measurements is to locate the antenna site on the map. Then a north–south line is drawn through the site. Following this, a protractor is used to draw the radial. (See Fig. 14-3 for a typical map with a radial.) The map should then be studied to find suitable measurement sites. As many potential measurement sites as practicable should be identified along each radial. Except in rural areas, many sites will prove to be unacceptable because of the proximity of conducting structures.

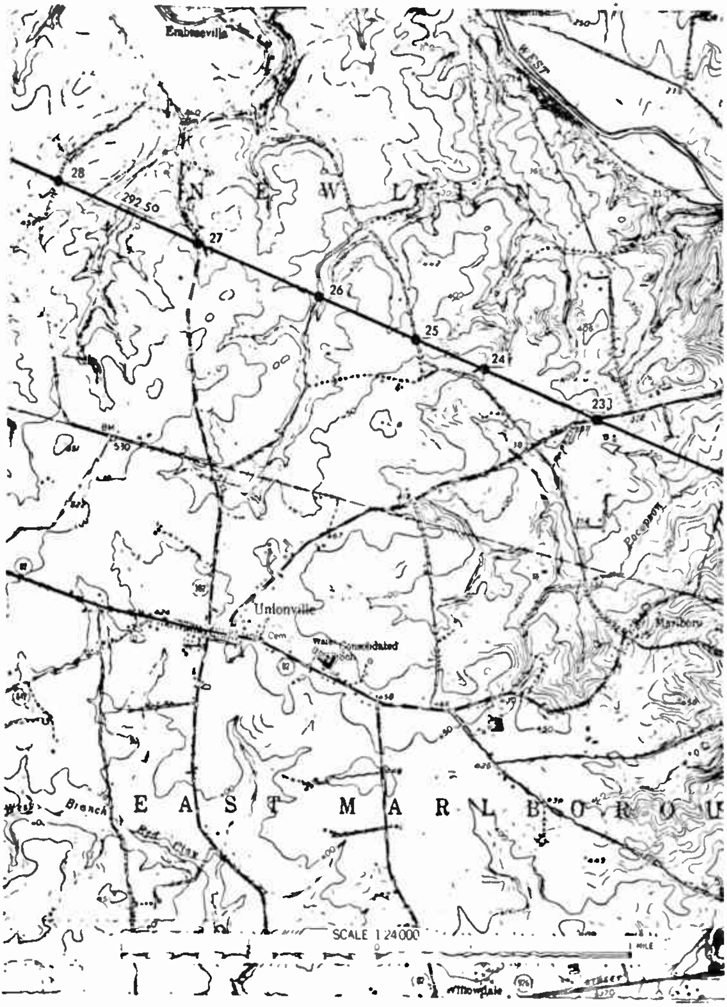


Fig. 14-3. Map with radial.

The measurement sites must not be too close to the antenna site. For a single-tower nondirectional antenna the nearest measurement site should be no closer than five to ten times the height of the tower. For a directional antenna the closest measurement site should be no closer than ten times the spacing between the most widely separated tower. The reason for this last condition is that a directional antenna is not much like a point source until one is a considerable distance from it.

When the first satisfactory measurement site is located and the measurement made, both the measurement and a description of the site should be entered in a log (Fig. 14-4). A column should be provided in the log for entering the radial distance from the array to the measurement site. This distance can be computed from the map and entered in the log later. The importance of proper identification of the measurement point cannot be overemphasized. Directional antennas cause enough problems without introducing others by making measurements that cannot be duplicated.

STATION: WXYZ
POWER: 5 kW
FREQ.: 1300 kHz
MODE: NONDIRECTINAL—TOWER
RADIAL: 40 TRUE
DATE: 3 – 13 – 76

MEAS. POINT	DISTANCE (MILES)	FIELD INTENSITY	TIME	COMMENTS
8	1.2	152.5	1300	IN FIELD, 100 FT DUE NORTH OF RT. 304.
9	1.7	148	1348	NE CORNER 8TH LANE, AND RT 4.
10	2.2	147	1420	BY LAKE—SEE PHOTO.

Fig. 14-4. Typical measurement-log entries.

The first series of radial measurements are made with the antenna operating in the nondirectional mode. If the antenna is to normally operate this way, these measurements will establish the efficiency. With a normally directional antenna, the nondirectional measurements will be of great value in evaluating its directional performance later on.

In general, the more measurements made along a radial, the better. When plenty of measurements are available, the data can be analyzed much more precisely. For a nondirectional antenna at least 25 measurements should be made on each radial. For a directional antenna 40–50 measurements should be made on each radial.

The number of radials on which measurements are made depends on the complexity of the pattern.

FINDING THE EFFECTIVE FIELD AT ONE MILE

After all of the measurements have been taken along a radial, it is necessary to analyze them to determine the effective field at one mile from the antenna and the effective ground conductivity. This is most conveniently done with the aid of charts published by the FCC. These charts appear in Part 73 of the FCC Rules, and a typical one is shown in Fig. 14-5.

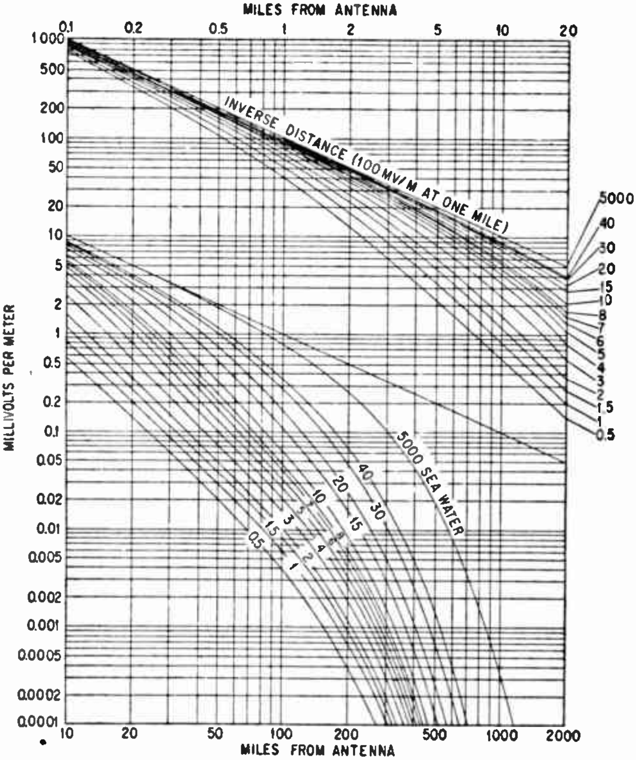


Fig. 14-5. Groundwave field intensity versus distance, 970–1030 kHz.

FCC Field-Intensity Charts

Before attempting to use the curves of Fig. 14-5, one must be thoroughly familiar with them and just what they are intended to show. The actual FCC curves show another family of curves that extended out to 2000 miles, but these are not used by the broadcast-station engineer in checking his antenna. Referring to Fig. 14-5, at the top of the graph is a straight line that shows how the signal would be attenuated over a perfectly conducting earth. Note that the field intensity at one mile is 100 mV/m, whereas at 2 miles it is 50 mV/m. This is just what we would expect, since over a perfectly conducting earth the signal varies inversely with distance. Now, looking at the curve labeled 10, we can find the attenuation of a signal over ground having a conductivity of 10 mmho/m. Note that in this case the field intensity at 2 miles from the antenna is only about 20 mV/m. The curve labeled 2 shows the attenuation of a signal over ground having a conductivity of 2 mmho/m, which is about the lowest conductivity found in the U.S. In this instance, the field intensity at 2 miles from the antenna would be about 17 mV/m.

The FCC curves are based on a field intensity at one mile from the antenna of 100 mV/m, but they can be easily scaled to handle other field intensities. For example, if our station produced an effective field of 500 mV/m at one mile, we could simply multiply all of the other points on the curves by 5.

Finding the Conductivity

Before we can use the FCC curves for a particular station, we must know the ground conductivity along each radial. It is important to realize that the conductivity might be different along different radials. We find the conductivity by plotting the measurements of field intensity on translucent graph paper with grid lines that exactly match the graph paper used in plotting the FCC graphs.*

Figure 14-6 shows a plot of measurements taken along a radial of 290° true from a station operating on 1460 kHz. Once plotted, our graph (Fig. 14-6) is placed on top of the

* This graph paper is available from Keuffel & Esser Co., 1766 N St., NW, Washington, DC 20036, as graph paper 99-1101 (orange lines) or 99-9901 (green lines).

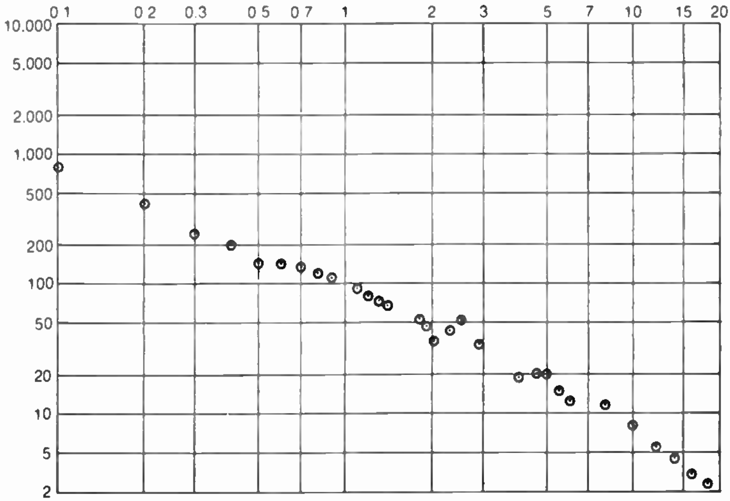


Fig. 14-6. Plot of measurements.

corresponding FCC graph (Fig. 14-5) so that the vertical grids of the graphs exactly coincide. Both graphs are then held up to the light so that the lines on the FCC graph show through. Then our top graph is slid up and down until its measurement points fall on one of the curves of the FCC graph. In our example, the measurement points fall on the 20 mmho/m conductivity curve of the FCC graph. When the two graphs are superimposed, the result looks like the composite graph shown in Fig. 14-7. Note that the grid lines of this graph are like those on the graph we plotted in Fig. 14-6, but the shape of the conductivity curve is like that on the FCC graph in Fig. 14-5.

Graphical Solution for Field

To form the graph of Fig. 14-7, we trace the 20 mmho/m curve on our graph of measurements. Also, we trace the straight inverse-distance line on our graph. Now we can find the effective or unattenuated inverse field at one mile from our station. Note that it is the point where the inverse-distance trace from the FCC graph crosses the one-mile ordinate on our graph. In the example of Fig. 14-7, the effective field at one mile from the antenna is 120 mV/m, even though the actual measured field was only about 108 mV/m.

The above procedure is extremely important in analyzing field-intensity measurements from a standard broadcast

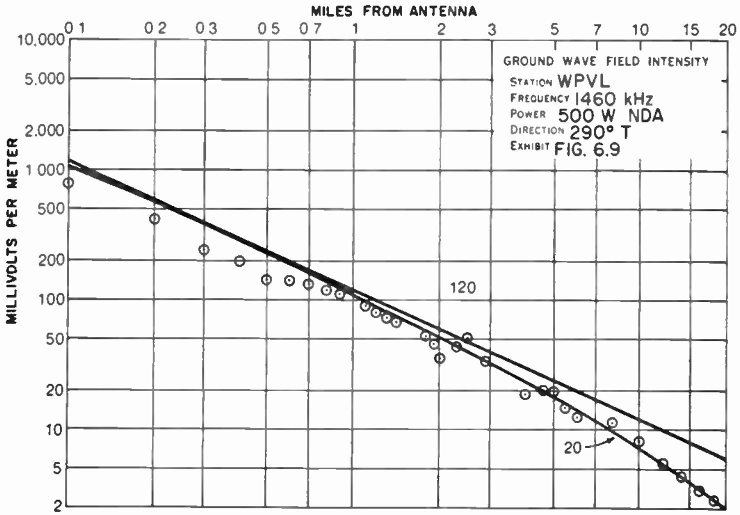


Fig. 14-7. Figures 14-5 and 14-6 superimposed.

station. Every broadcast engineer should be thoroughly familiar with it. For this reason the step-by-step procedure is given below.

1. Plot measurements of field intensity versus distance on the proper type of graph paper.
2. Place the graph on top of the proper FCC graph so that the vertical grid lines of the two graphs coincide exactly.
3. Hold the two graphs up to the light so that the coordinates of both graphs can be read.
4. Slide the top graph up and down until the measurement points on it fall on one of the curves of the FCC graph. Keep the two graph sheets in this position.
5. Trace the proper conductivity curve of the FCC graph onto the graph that has been plotted so that the curve passes through the measurement points. This yields the ground conductivity along the radial.
6. Trace the inverse-field curve from the FCC graph onto the graph that has been plotted. The point where the inverse-field line crosses the one-mile ordinate on the top graph gives the effective field at one mile from the antenna.

It is rare that all of the measurements that have been made along a radial will fall exactly on one of the conductivity curves published by the FCC. One method of making an approximate fit is to have as many of the measurement values appear above the curve as fall below it.

It might appear superficially that the effective field at one mile from the antenna is a rather arbitrary figure on which to base allocations, but inspection of the method of deriving the figure will show that it is actually based on a series of measurements and tells much more about the performance of an antenna than any single measurement possibly could.

Curve-Fitting

When measurements are made on a nondirectional antenna or a directional antenna operating in the nondirectional mode, the measurement points usually fit quite nicely on one of the conductivity curves on the FCC graphs. Occasionally, there is an abrupt change in ground conductivity along one or more radials. Then part of the measured data will fall on one of the conductivity curves, and the rest will fall on another curve. Such a situation is shown in Fig. 14-8. Here the measurements out to about 5 miles fall on the 10 mmho/m

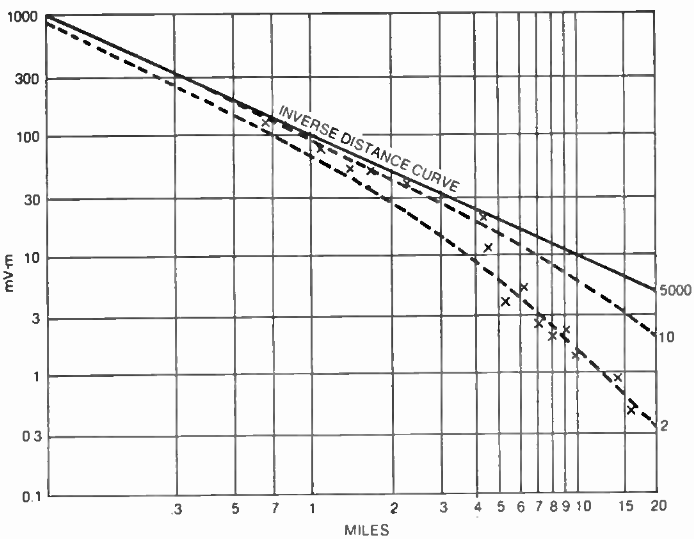


Fig. 14-8. Break in ground conductivity of similes.

curve. Beyond that the points fall on the curve corresponding to 2 mmho/m. In this case, there is an abrupt change at about 5 miles from the antenna, where the land becomes sandy and dry.

Usually directional-antenna measurements do not fit the conductivity curves as well as nondirectional-antenna measurements, particularly when made close to the antenna. One reason for this is that the directional antenna doesn't look as much like a point source as a nondirectional antenna does. By looking at the plot of field intensity versus distance along a radial of a directional antenna, we will usually see that the data behaves rather wildly when the measurement points are too close to the array. Figure 14-9 shows a plot of field intensity versus distance along one radial of a 2-tower array. Note that the field intensity actually increases with distance at certain points. In such a case it is better to rely on measurements taken at greater distances. Unfortunately, the radial may lie along a rather deep null in the pattern, and there may not be much signal strength to start with.

There are other reasons for data skewing along the radial of a directional antenna, and sometimes these can't be pinned down; but the data is simply hard to place on any particular

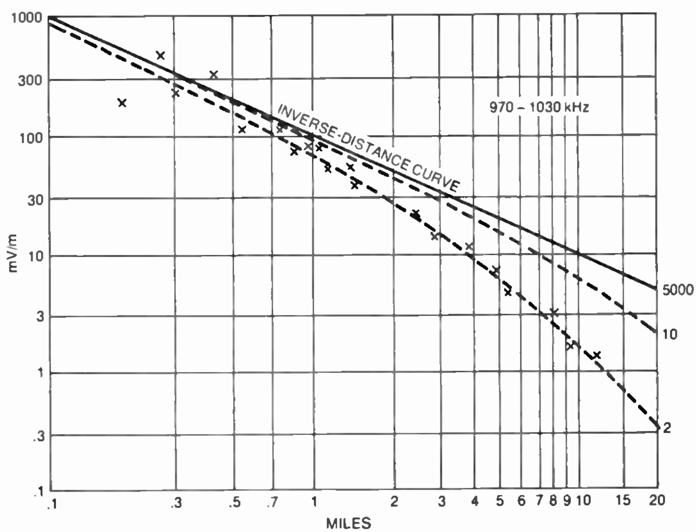


Fig. 14-9. Data spread near antenna.

conductivity curve. In such cases, it is sometimes helpful to take the ratio of the directional and nondirectional measurements taken at the same points. We then try to plot the *ratios* of the measurements to see if we can make them fit one of the FCC curves so that we can find the inverse field at one mile along the radial. If the ratios fall on one of the curves, the FCC will probably accept our data. When the ratios don't fit, we can try other things, such as taking the ratio of each measurement to the average ratio.

This problem is most severe on radials in deep nulls where the signal must be held to a low value to protect other stations. There isn't much field intensity along such a radial in the first place, so the measurements are much more susceptible to such things as reflections from conducting objects.

The interpretation of field-intensity measurements that do not fall into the normal pattern is more of an art than a science. The FCC is often helpful in this, and there are many consultants who specialize in directional-antenna work.

CHECKING THE PATTERN

The purpose of radial measurements is to establish or check an antenna pattern. The procedure used in checking a pattern is usually called a *proof of performance*. The initial proof of performance is made before the station license is issued. A directional-antenna system is built on the authorization of a *construction permit*, which is based on a theoretical pattern that is computed in accordance with the procedures given in earlier chapters. Then a complete set of measurements is made to assure that the actual pattern is within the prescribed limits. If it is, the station license is issued.

Until 1971 each standard broadcast station with a directional antenna had two antenna patterns. The first was the theoretical pattern on which the construction permit was based. The second was a revised pattern based on actual measurements. Figure 14-10 shows a typical pattern. The solid line represents the theoretical pattern. In practice, it proved impossible to attain this pattern, so the limits were relaxed as shown by the dotted lines. The dotted lines represent the

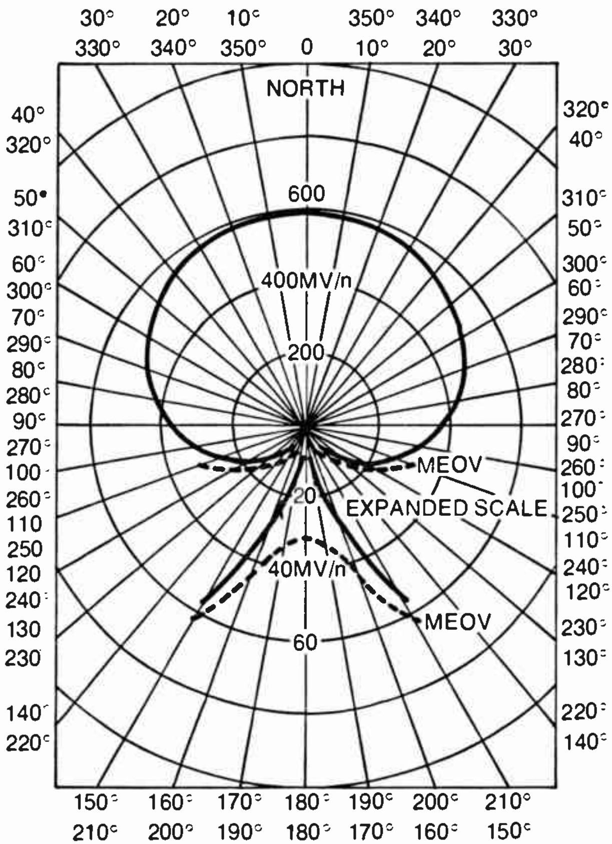


Fig. 14-10. Pattern with MEOV.

maximum expected operating value or MEOV, pattern. Since 1971 all new installations and those that have had major modifications have only one antenna pattern, which is referred to as the *standard pattern*.

In addition to the complete proof of performance that is required when a station is installed and when major modifications are made to the antenna system, *partial* or *skeleton proofs* are required under other conditions. When a station makes a request of the FCC to operate with parameters at variance from their licensed values, the Commission may require a skeleton proof of performance. A skeleton proof is also required when a station using remote control applies for license renewal. A skeleton proof consists of three or four field-intensity measurements made on each radial specified in the original complete proof.

When a station applies for permission to use third-class-licensed operators for routine duties, it must submit a partial proof, which consists of at least 10 field intensity measurements made on each radial.

The number of radials on which measurements must be made depends on the complexity of the pattern. Eight radials are required to establish the nondirectional characteristics of an antenna, and if the pattern has many nulls and major lobes, many more may be required. The FCC considers three radials sufficient to establish a major lobe.

Once the radial measurements have been made and the effective field at one mile determined for each radial, the pattern may be plotted. It is usually a good idea to plot the pattern to the same scale used in the original pattern so that a quick visual check can be made to see if there are any drastic changes.

Suppose that a series of radial measurements have been made on 12 radials, and the effective field at one mile along each radial has been found to be as shown in Fig. 14-11. The pattern can then be plotted as shown in Fig. 14-11. Let us assume that the pattern is within the prescribed limits.

The next step is to check the *size*, or the *rms value*, of the pattern. (The rms or root-mean-square value should not be confused with the rss root-sum-square value. The rss figure is used in computing the standard pattern of the station.) The rms value of an antenna pattern is the radius of a circle, expressed in millivolts per meter, that has the same area as the pattern.

There are several ways in which the rms value of an antenna pattern can be calculated. One way (described in Chapter 7) is to determine the area enclosed by the pattern on a graph and calculate the radius of a circle of the same area. Another method is to square the effective field at one mile on each radial, find the average of the squares, and take the square root of the average. Using the example of Fig. 14-11, we would find the rms field as follows:

Effective Field	*(Effective Field)²*
620	384 400
600	360 000
518	268 324
372	138 384
164	26 896
62	3 844
17	289
59	3 481
203	41 209
354	125 316
422	178 084
600	360 000
	1 890 257

$$E_{ave}{}^2 = \frac{1{,}890{,}257}{12} = 157.521 \quad E_{rms} = \sqrt{157.521} = 397 \text{ mV/m}$$

This gives the rms field as 397 mV/m, which is only approximate. This method is only accurate when we have many radials, say, one every 5° of azimuth.

The rms value of the pattern must be checked against the original value or any previous value that has been accepted by the FCC. Any substantial change in the rms value means that there has been a change in the antenna system and that it might no longer meet the license requirements. The usual change is that the rms value becomes smaller, which often means that the losses associated with the antenna or feeder system have increased. A frequent cause is deterioration of the ground system. The rms value will also decrease if the transmitter output power is lower than it should be. The common-point impedance and current should be checked for this.

Sometimes, after extensive repairs to a directional antenna, the rms value of the pattern will increase. This is most apt to happen if the feeder system has been updated or the ground system has been replaced.

MONITORING POINTS

The station license of every standard broadcast station that uses a directional antenna specifies a number of locations

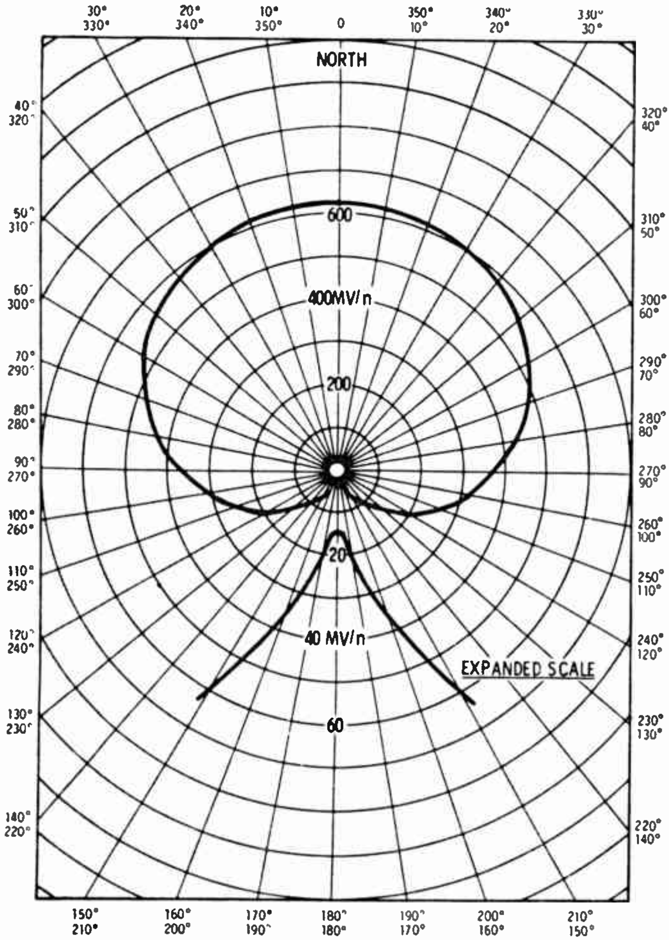


Fig. 14-11. New standard radiation pattern.

where measurements must be made at regular intervals to establish that the pattern has not changed. These are called *monitoring points* and are proposed in the original license application. The number of monitoring points depends on the complexity of the pattern. The field intensity must be measured at each monitoring point and the values entered in the station's maintenance log once every 30 days unless the station license requires a shorter interval between measurements.

In general, there are three things that can cause the measurement at a monitoring point to deviate from the value specified in the station license:

1. *A change in the operation of the antenna system.* This is sometimes the reason for making the measurement. Care should be taken to insure that the change is actually in the antenna system before any corrective action is taken.
2. *A change in the environment of the monitoring point.* The usual change is new construction, such as power lines, water towers, or apartment buildings. Usually such construction will make the monitoring point useless, and permission must be sought to use an alternate monitoring point.
3. *A change in the conductivity of the soil.* This is most likely to occur in the northern part of the U.S.

The FCC Rules require that each monitoring point be properly identified. The identification should include a photo of the field-intensity meter in each location and a map showing the most accessible route to each location. This requirement may seem trivial, but often when a new engineer takes over at an old station, one of his most serious problems is finding the monitoring points.

Good monitoring points are even more important than good points for running a radial. With a radial, many measurements are made, and if one of them happens to be made from a site that isn't really suitable, the effect will be averaged out when we compute the effective field. With a monitoring point, we normally have only one point on a radial, and if the site isn't suitable, a partial proof of performance may be required to establish a new monitoring point. The evaluation of measurement points was described in connection with radials.

Although most of the measurements required by the FCC can be made at monitoring points or along radials, sometimes a great deal of information can be gathered by making measurements along an arc drawn through a monitoring point (Fig. 14-12). These measurements can be used to verify the

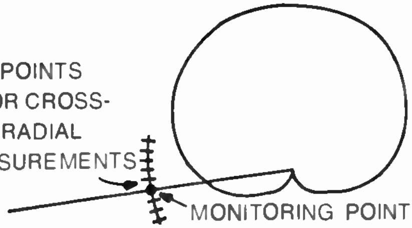


Fig. 14-12. Selecting points for cross-radial measurements.

shape of a null in the pattern, to find out whether the position of a null has shifted, or sometimes even to find a null in the pattern that wasn't supposed to be there.

The best way to make these cross-radial measurements is to carefully locate a few suitable measurement sites on a map and then make both directional and nondirectional measurements at each point. The ratios of the nondirectional and directional measurements are found and then plotted (Fig. 14-13). The ratio points will define the shape of the pattern. Any shift in the pattern can thus be easily detected. When more than one monitoring-point measurement is out of limits, the cause may be a shift in the pattern. One possible source of a shift is changes in the phase-monitoring system. If the monitoring system changes slowly, the engineer may think that the change is in the antenna system and adjust the phasor accordingly. As a result he will be rotating a perfectly good pattern to compensate for changes in his monitoring system.

SEASONAL VARIATIONS

One of the most perplexing effects in directional antennas in the northern part of the country is a change in signal

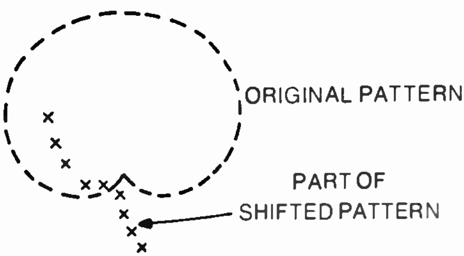


Fig. 14-13. Cross-radial measurements showing that pattern has rotated.

strength at monitoring points with changing temperature. Most engineers agree that the seasonal change in monitoring-point measurements has something to do with soil conductivity, but there is no general agreement as to the actual mechanism involved.

In many stations the monitoring-point measurements made in the winter are considerably higher than those taken at other times. It is almost axiomatic that at northern latitudes the original proof of performance should be made in the winter. That way the monitoring-point measurements are not likely to exceed the licensed values.

When a seasonal change in monitoring-point measurements is noted, one must establish whether or not there has been a change in radiation at the antenna. The first thing to look for is any change in current ratios or phases. If the ratios and phases haven't changed, the monitoring-point change is probably due to a change in soil conductivity. If the field-intensity change is downward there is no serious problem. If the change is upward so that the licensed value for the monitoring-point measurement is exceeded, the station is operating in violation of the FCC Rules. The first thing to do is to notify the Commission and request permission to operate at

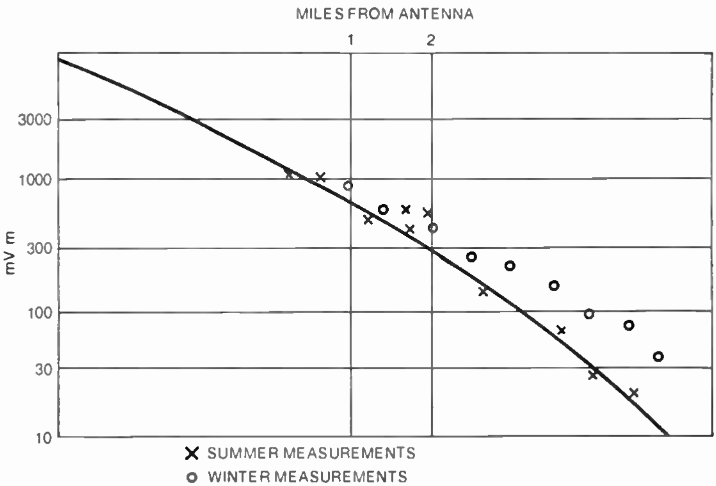


Fig. 14-14. Seasonal change in ground conductivity as determined by FCC curves.

variance with the licensed value. At the same time explain what you are doing to correct the problem. A field change due to soil-conductivity effects can be verified by making radial measurements and plotting them against the FCC conductivity curves, just as we did to find the effective field at one mile. If the field change is found to be caused by a change in soil conductivity, the new measurements will fall on a different conductivity curve, as shown in Fig. 14-14.

The curves of Fig. 14-14 show that the change is greatest at the points farthest from the antenna. Hence, the closer a monitoring point is to the antenna, the less subject it is to conductivity effects. This means that monitoring points should be as close to the station as is consistent with reliable measurements. Data of the type shown in Fig. 14-14 may be submitted to the FCC with a request for a change in monitoring-point values.

Undoubtedly, as time passes, more will be learned about seasonal variations in the field, and better solutions to the problem will be discovered.

Chapter 15
Parasitic Reradiation

The reradiation of radio energy by structures is not a new problem. In 1922, the National Bureau of Standards investigated the influence of the Washington Monument on the fields from distant stations. They found that the fields were perturbed enough to influence the accuracy of radio direction finders used in the vicinity of the monument. In 1924, Henry C. Forbes of the University of Minnesota reradiated a signal from a local broadcast station. Forbes inserted a telegraph key between his antenna and ground and transmitted CW signals a distance of about 3 miles. using only the reradiated signal from the broadcast station over a mile away.

The reradiation problem manifests itself in broadcasting as a change of the antenna radiation pattern with the erection nearby of structures such as power lines, water towers, buildings, and bridges. Often, the station has no control over the structures that are erected in the vicinity of its antenna, and the first sign that a newly erected structure has influenced the antenna pattern is when monitoring-point measurements become abnormal.

GENESIS OF THE PROBLEM

When a broadcast station is constructed, a site is selected away from any large conducting structures that might influence the radiation pattern. The site is usually in a rural area, often on high ground. To provide a good signal over the city of license, the station is often situated at the outer edge of the suburban area. Unfortunately, as the community grows, this is the area where growth is most likely to take place. New power lines are run into the area, and if the site is on high ground, it is a likely place for a water tower to be built. Too, this is the type of location often used by light industry, trucking companies, and telephone-switching centers, all of which are making increasing use of two-way radio and may erect towers to support their antennas.

In short, practically every standard broadcast station is subject to the problems that may result from the erection of new conducting structures. If the potentially offending structure is intended to support a two-way radio, microwave, or broadcast antenna, recourse to the FCC is possible. Sometimes the expense of solving the problem may be borne, at least in part, by the owner of the structure. In other cases, the station may have to bear the expense. In almost all cases, the broadcast engineer or consultant must do the necessary engineering to solve the problem.

There are two general types of problems that result from parasitic reradiation. Some structures reradiate enough energy to actually change the radiation pattern of an antenna. These are usually structures that are over 1/8 wavelength in height at the broadcast frequency. In general, the closer the structure is to the antenna site, the more it influences the antenna pattern.

A second and lesser problem results when a smaller conducting structure disturbs the field locally to a sufficient degree to make field-intensity measurements at one or more monitoring points fall outside of their limits, but doesn't significantly affect the overall radiation pattern. This problem can be solved by getting permission from the FCC to use another monitoring point on the same radial. This shows the wisdom of providing alternate monitoring points whenever an application is filed with the FCC for a new station or for modifications to an existing one.

When potential reradiation is spotted early enough, the engineer can contact the responsible parties and possibly work out a solution before the problem actually exists. In other cases, construction is started in an area where it isn't likely to be noticed by the engineer before it causes a problem. The only thing that can be done then is to seek FCC permission for operating with parameters at variance and to start looking for the offending structure.

Most reradiation problems can be solved by detuning the offending structure. This approach is facilitated if cordial relations are maintained with the owner of the structure since it involves alterations to his property.

LOCATING PARASITIC STRUCTURES

Before we can take any steps to minimize parasitic reradiation, we must first find the offending structure. The obvious tool to use for the job is the field-intensity meter. There is only one problem with this: The field-intensity meter is to be used to locate a comparatively weak field from a parasitic radiator in the presence of the strong field of the station. We can make such a measurement by taking advantage of the null of the meter's loop antenna at 90° from the direction of maximum pickup. Thus, if we find a place where a line to the station antenna and a line to the suspected structure intersect at a 90° angle, as in Fig. 15-1, we can measure the radiation from the structure with minimum interference from the station signal.

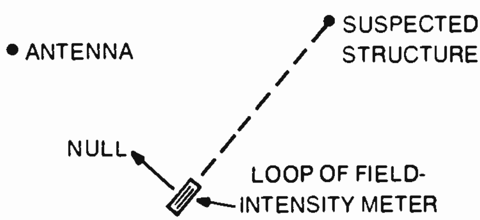


Fig. 15-1. Minimizing signal from antenna.

There is a principle of geometry that we can use to find suitable measuring points (see Fig. 15-2). If the station antenna and the suspected structure are diametrically opposite each other on a circle, at any point on the circle the angle between lines to the ends of the diameter will be a right angle. Thus the angles at points *A*, *B*, and *C* in Fig. 15-2 are right angles. Therefore, if we locate our field-intensity meter at any point on the circle, when the direction of maximum pickup is toward the suspected structure there will be a null in the direction of the station. Thus interference from the direct signal of the station will be minimized.

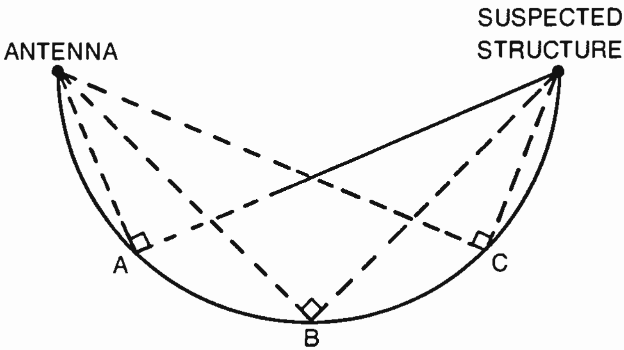


Fig. 15-2. Measurement points for verifying parasitic reradiation.

To use this principle, the engineer should get a good map of the area and draw a circle with the station antenna and suspected structure at opposite ends of a diameter. Then suitable measurement points can be picked out on the circumference of the circle. Measurements made at each of these points will determine whether or not a suspected structure is actually reradiating and, if so, the intensity of the radiation. Naturally, reradiation is more serious in patterns with deep nulls to protect other stations.

Not every tall structure nearby will affect the pattern of a broadcast station. The effect on the pattern depends on the height of the structure. In one case, a microwave tower was erected, one section at a time, near a 4-tower directional-antenna system. As construction progressed, the current in one tower of the antenna system began to drop as each new section was added to the microwave tower. The current continued to drop until it nearly reached zero. But, as more sections were added to the microwave tower, the current began to rise. When the microwave tower reached its final height, all of the tower currents of the broadcast antenna, as well as its pattern, were normal.

EVALUATING PARASITIC RERADIATION

In cases of possible reradiation, all suspected structures should be carefully marked on a map. Then measurements should be made at several points selected in accordance with the principle illustrated in Fig. 15-2. If there is a strong field from the structure and the field intensity has changed at more than one monitoring point, it is reasonably certain that the radiation pattern of the station has changed, and the only recourse is to attempt to detune the structure. If the reradiation is comparatively weak and the offending structure is close to a monitoring point, it may only be necessary to establish a new monitoring point on the same radial. To prove that this is the case, it may be necessary to run measurements along the radial and compute the inverse field at one mile and compare it with the original computations.

A parasitic radiator, like a regular antenna element, has both an induction and a radiation field. As in a regular antenna element, the induction field varies inversely with the square of the distance from the element, whereas the radiation field varies inversely with the distance. The induction field is apt to upset the field intensity at a nearby monitoring point, but it dies out so rapidly with distance that it does not disturb the pattern of the station antenna. The radiation field is the one that causes serious problems in the pattern.

It is possible for stations that have nondirectional or comparatively simple patterns to have reradiation problems that do not show up in ordinary field-intensity measurements. The first clue to problems of this sort is usually complaints of poor reception. If a rash of poor-reception complaints is received, the location of each complainant should be pinpointed on a map. If several complaints of this type come from one area, the area should be searched for a reradiating structure. Often complaints of this type result from a structure casting a "shadow" over a certain area.

AVOIDING PROBLEMS OF RERADIATION

Although usually the broadcast engineer has no control over the location or configuration of structures that may cause reradiation problems, there are some cases where he has some control and can avoid reradiation problems. For example, if the licensee of a standard broadcast station intends to erect a tower for a television station or an STL antenna, the broadcast engineer may be able to influence some of the specifications. Often, many different stations have their antenna facilities in the same general area on the outskirts of a city. Usually these stations will cooperate to forestall parasitic problems. If a structure is being built with the help of federal funds, it may be possible to bring pressure to bear on the responsible government agency to avoid parasitic problems.

As a rule, if a structure doesn't receive much energy, it can't reradiate much. If possible, the offending structure should be located in an area of minimum signal from the broadcast antenna. If it is possible to build a structure in short sections with insulators, this will minimize the problem. Similarly, guy wires should be broken up with insulators.

MECHANISM OF PARASITIC RERADIATION

Most of the antenna elements that we have considered are fed with transmission lines but a line isn't always necessary. An antenna element may be fed directly by the electromagnetic field of another antenna. In fact, many antennas used at higher frequencies use parasitic elements that have no connection to the transmission line.

The phenomenon can be understood from Fig. 15-3, which shows 1/2 wavelength of wire in the field of a broadcast antenna. For convenience, the wire is oriented vertically and we neglect the influence of the ground. Inasmuch as the wire is parallel to the electric component of the field, it will tend to short the electric field, and current will flow in the wire. That

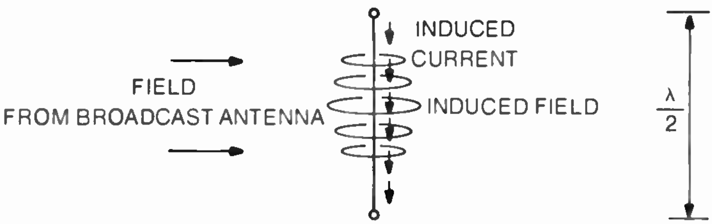


Fig. 15-3. Parasitic excitation of a wire

is, the wire will be excited just as though it were connected to a source of energy through a transmission line. It will carry a current at the same frequency as the incident field and will set up an electromagnetic field of its own. The two fields will combine vectorially, and the resultant field will depend on the magnitude and phase of the field from the wire. These, in turn will depend on the magnitude and phase of the current induced in the wire.

The effect of the incident field on the wire will be the same as if there were a series of small generators distributed throughout its length as shown in Fig. 15-4A. The current distribution won't be nearly sinusoidal, as it would with one generator, but will have a bell-shaped distribution curve (Fig. 15-4B). The current will be zero at the ends of the wire because there is no place for it to flow. Conversely, the *voltage* will be *maximum* at the ends. The maximum current will be at the center. We would have almost exactly the same thing if we had a 1/4-wavelength wire grounded at the bottom as in Fig. 15-4C. Here the current would be zero at the top and maximum at the grounding point. From these examples we can see that we are apt to have reradiation problems from grounded structures that are approximately a 1/4 wavelength in height and from

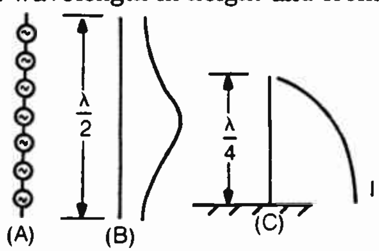


Fig. 15-4. Parasitic structures.

ungrounded structures that are about 1/2 wavelength in height.

In only a few cases will a structure be an exact fraction of a wavelength in height. From our study of antenna impedances we know that an element that is not exactly resonant has some reactance; thus the current will not be in phase with the voltage. In reradiating structures this means simply that the phase of the reradiated field is not necessarily either exactly in phase or exactly out of phase with the incident field.

A hint of how we might cope with a reradiating structure is given in Fig. 15-5. Here we have a grounded quarter-wave structure. As pointed out above, the current will be maximum at the bottom and zero at the top. Suppose we could insulate the base of the structure from the ground as in Fig. 15-5B. Now the current will have to be zero at the bottom as well as at the top because there is no place for it to flow. (This isn't strictly accurate, because there would be a capacitive path across the insulator; but it is accurate enough to illustrate the principle.)

The parasitic structure of Fig. 15-5B would effectively be a short antenna, in terms of wavelength. This means that its radiation resistance and hence its efficiency would be very low. Such a modified structure is a step in the right direction because it is not an efficient radiator. Another property of a short antenna is that its impedance is highly reactive. If we wanted such a structure to radiate, we would try to tune out its reactance. Naturally we wouldn't do this to a structure we didn't want to radiate.

It isn't feasible to insulate a structure such as a water tower from the ground. There are, however, ways that we can make it behave electrically as if it were insulated from ground.

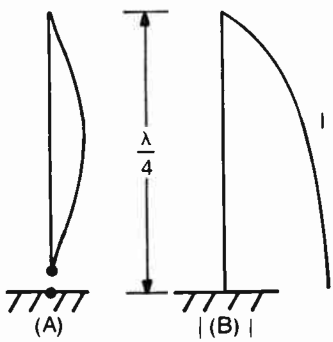


Fig. 15-5. Suppressing radiation by insulating a structure from the ground.

MINIMIZING PARASITIC RERADIATION

The most troublesome structure is a vertical grounded conducting structure that is about 1/4 wavelength high. This structure acts like a grounded quarter-wave antenna and reradiates the energy it intercepts quite efficiently. To see how we might minimize reradiation from such a structure, it helps to look at the way we minimize radiation from a transmission line.

In a transmission line, charges are in motion, and they tend to make all other charges in the universe move in the same way. But the currents in the two conductors of a transmission line are equal and flow in opposite directions. Thus any effect that the current in one conductor might have on a distant charge is canceled by the opposite effect of the current in the other conductor. Now, if we can add another conductor to a reradiating structure and cause the current in it to flow in the opposite dirction of the current in the reradiating structure, there will be little or no radiation from the combination.

Figure 15-6A shows a quarter-wave grounded vertical structure. An incident field causes current to flow in it just as in a quarter-wave vertical antenna. To this structure we can add other conductors in an outrigger configuration (Fig. 15-6B). We now need to make the current in the outrigger equal to that in the structure and opposite in direction. If the conductors are completely insulated from each other, as in Fig. 15-6B, the outrigger will be capacitively coupled to the structure, and the whole affair will reradiate just as much as the structure alone. If we connect the outrigger to the structure at the base, as in Fig. 15-6C, we will have the same situation. The assembly will behave electrically like a single conductor and will reradiate.

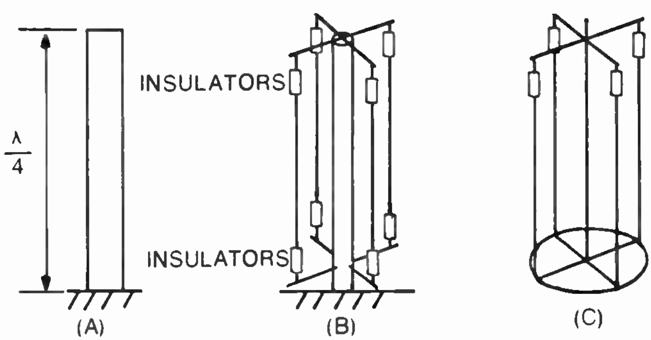


Fig. 15-6. Quarter-wave structure and outriggers.

If we connect the outrigger to the structure at the top (Fig. 15-7A), we have an entirely different situation. Electrically, the structure now looks like a quarter-wave coaxial transmission line shorted at one end (Fig. 15-7B). The structure may be thought of as the inner conductor and the outrigger as the outer conductor. If one end of a quarter-wave section is shorted, the other end looks electrically like an open circuit. Thus, inasmuch as the structure (inner conductor) is connected to ground, the outrigger will have a very high impedance to ground.

For current to flow into one conductor of a transmission line, an equal current must flow out of the other conductor. If current tries to flow between the structure and ground in Fig. 15-7B, an equal current should flow between the outrigger and ground. But this is impossible because there is a very high impedance between the outrigger and ground. What we have done, in effect, is to make the structure a quarter-wave conductor isolated from ground, which is not a very efficient radiator.

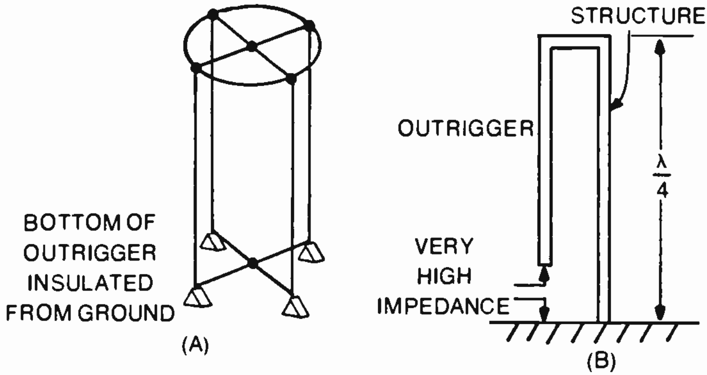


Fig. 15-7. Making an open quarter-wave stub with an outrigger.

The above case is oversimplified in that the structure to be detuned happened to be exactly a 1/4 wavelength high. Most offending structures aren't this agreeable. Although the most troublesome structures are those that are close to odd multiples of 1/4 wavelength, only rarely is the height an exact multiple of 1/4 wavelength. In such cases, we can use a simple short circuit to connect the outrigger to the structure, but we must use some reactance. The added reactance effectively lengthens or shortens the structure to the point where it behaves like a shorted quarter-wave line. Usually a tapped coil and capacitor, mounted in a weatherproof box as shown in Fig. 15-8, are used for the purpose. The coil and capacitor can be connected either in series or in parallel, and the position of the tap on the coil can be varied to provide a wide latitude of adjustment.

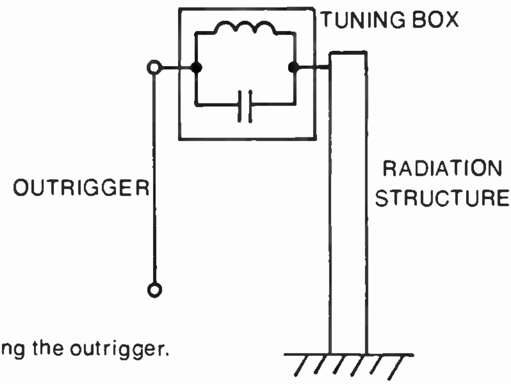


Fig. 15-8. Tuning the outrigger.

The structure and outrigger will not form a perfect transmission line, and the exact velocity of propagation will not be known accurately. Hence it is usually necessary to tune the network experimentally to minimize reradiation. The easiest way to do this is to have one man with a field-intensity meter stationed at a measurement point established as in Fig. 15-5, and another man at the structure making the adjustment. Two-way radio sets can be used to relay the field-intensity measurements to the man making the adjustment at the structure.

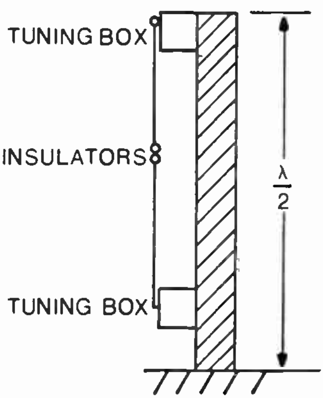
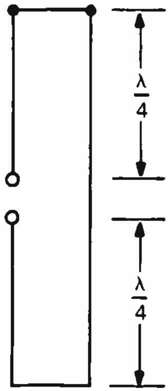


Fig. 15-9. Detuning a tall structure by effectively breaking it into sections.

Fig. 15-10. Equivalent circuit of Fig. 15-9.



Tall Structures

Often problems are encountered with structures that are much taller than 1/4 wavelength. The best approach to detuning them is to effectively break them up into shorter sections (Fig. 15-9). Each of the sections is effectively isolated from the others, and no substantial current flows in the complete arrangement.

Figure 15-10 shows the equivalent electrical circuit of the arrangement of Fig. 15-9. It consists of two shorted quarter-wave sections of transmission line. The impedance seen looking into the open end of the line is very high, so very little current flows into it. This arrangement reduces the total current flowing in the structure to the point where reradiation is minimal.

Water Towers

Water towers are common sources of parasitic reradiation. These towers come in various configurations but are always grounded. Figure 15-11 shows a water tower with an outrigger consisting of six wires. The arrangement is tuned by means of a coil and capacitor located at the top of the tower.

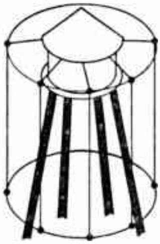


Fig. 15-11. Detuning a water tower.

Chapter 16

AM Antenna Maintenance

The discussions in this chapter are centered around the directional antenna because its maintenance is much more of a problem than that of a single tower. Furthermore, most of the procedures involving directional towers, networks, and feeder systems apply in a limited sense to nondirectional antennas as well.

There are four aspects to a good maintenance program:

1. *Design considerations.* The design of an antenna system definitely affects its ease of maintenance. Some factors that make an array hard to maintain may be inherent in the station's allocation and hence unavoidable; others can be avoided either during the original design or when a modification is made to the system.
2. *Tools.* These include instruments for measurements and such items as complete diagrams which can give a great deal of useful information about the system.
3. *Maintenance procedures.* These should be established and followed on a regular basis.
4. *Troubleshooting procedures.* It isn't always possible to avoid troubles, even with the most effective maintenance system. It is necessary to have established procedures that will quickly lead to the root of a problem.

DESIGN CONSIDERATIONS

Directional-antenna systems called *critical arrays* are antennas in which the pattern has deep nulls—a signal reduction of 20 to 1 or more—to protect the service areas of cochannel or adjacent-channel stations. In such systems it is necessary to hold very close tolerances on current ratios and phase angles. Limits of ±1% for current ratios and ±2° for phase angles are common. Even a small change in the value of one or more of the components of the feeder system of such an antenna will throw it out of limits. This isn't necessarily the result of poor design. Use of such a tight pattern may be the only way that the station can be licensed.

There are other design factors that make an array unstable or hard to maintain and can be avoided. Chief among these are the following:

1. Improper grounds or an inadequate ground system
2. Networks that match impedances that differ by more than 10 to 1
3. Inadequate sampling and monitoring systems

Each of these items has been discussed in earlier chapters in some detail. Whenever they are encountered in an array, they should be corrected at the earliest opportunity.

It may seem superfluous to say that a system must be properly installed in the first place if it is to be kept operating properly with a minimum expenditure of time and money. Nevertheless, there are many directional-antenna systems that have seemingly insurmountable maintenance problems that can be corrected by merely doing the things that should have been done when the system was installed. Many broadcast-station licensees have found that the cost of getting the station on the air exceeded their estimates, with the result that the final phases of installation were carried out on a very limited budget. The best the engineer can do is to identify the troublespots and make plans for correcting them at the earliest opportunity.

SYSTEM DIAGRAMS

The first tool the engineer needs to develop an effective maintenance program is a set of accurate diagrams of his system. Depending on the complexity of the system, this may be either one large diagram or a collection of smaller diagrams. A typical diagram of a 2-tower system is shown in Fig. 16-1. Note the details that are shown. Starting at the towers, the driving-point impedances, current values, phase angles, and radiated powers shown.

Proceeding back toward the transmitter, the parameters of the base-impedance matching networks are given in detail. The reactance of each coil and capacitor at the operating frequency is specified and the positions of the taps on the various coils are given. The length of each transmission line in electrical degrees is stated, and full details of the phase shift and power-dividing networks are given.

Complete information as in the diagram of Fig. 16-1 permits the engineer to check for changes in the system whenever any of the measurements indicate that something is wrong. With the details given, each component can be checked with an impedence bridge to assure that its value hasn't changed. The notation of the tap positions of the various coils in the feeder system lets an engineer put all of the taps back where he found them if his adjustments do not produce the expected results.

At each of the points labeled with a number in a circle in Fig. 16-1, a jack is provided for measuring current. If an operating bridge is available, the impedance at each of these points can be measured. The impedances are measured not only at the carrier frequency but at several frequencies on either side of the carrier. These measurements assure that there is no fault tending to limit the bandwidth of the array.

The importance of a very complete diagram as a maintenance tool cannot be overemphasized. Periodic measurements at the various points can be compared with earlier measurements to detect trends that indicate changing values of components. Then defective components can be replaced before a catastrophic failure occurs and shuts the station down.

A detailed drawing can also disclose design weaknesses in the system that may lead to instability. For example, a network that transforms impedances over a range of greater

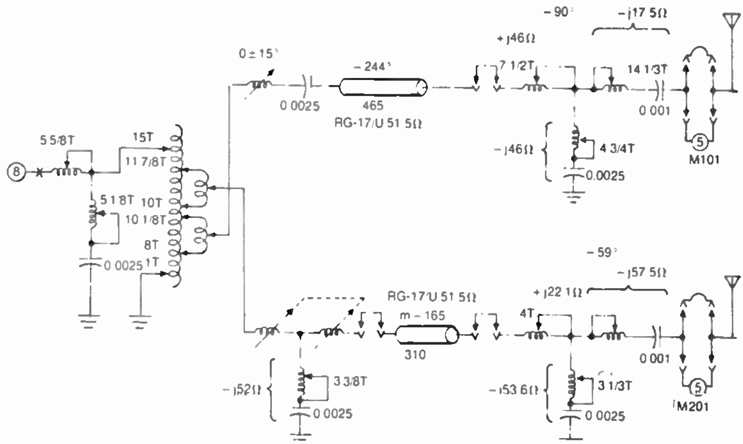


Fig. 16-1. Detailed diagram of feeder system. (Courtesy Smith Electronics, Inc.)

than 10 to 1 can be spotted easily. Such a network has high losses, narrow bandwidth, and a tendency toward instability.

To avoid complicating the diagram of Fig. 16-1, the monitoring system is not shown. But an accurate diagram of the monitoring system is fully as important as the system diagram. After all, the monitoring system is one of the means we use to find trouble in the antenna system. It must be maintained fully as well as, if not better than, the antenna system itself.

As with the feeder system, complete details of the monitoring system are shown on its diagram, including cable lengths and the open- and short-circuit impedances of all of the sampling lines. If for any reason it is suspected that the monitoring system has changed, measurements may be made and compared with the values on the diagram.

MAINTENANCE RECORDS

Next to the diagrams the most valuable maintenance tool is a very complete set of maintenance records. These should include a list of things to be checked periodically, as well as a record of what was observed. Although the FCC Rules require that a maintenance log be kept, additional maintenance records should be kept, and many more measurements should be made than are required by the Rules. The better the maintenance records, the easier it is to keep the array functioning properly. It is always a good idea to log weather conditions daily, including temperature, humidity, and storms. This information can often be correlated with a serious change in some measured value to give an indication of the cause of the abnormality. This is particularly true of thunderstorms and windstorms.

Whenever any quantity that is measured periodically is seen to change value unexpectedly, it is a good idea to plot the measurements against time, with notations of weather conditions added. Such a plot is shown in Fig. 16-2. Here we can see that although the base current to a tower probably varied too much at all times, the measurements clustered about a new value after a particularly severe thunderstorm. This would cause us to look for damage that might have been caused by a lightning surge.

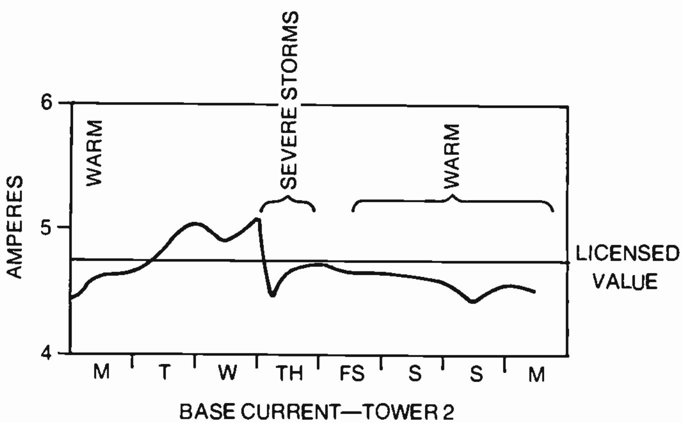


Fig. 16-2. Plot of meter variations.

MAINTENANCE SCHEDULE

Almost every feature, either mechanical or electrical, of a directional-antenna system should be checked at regular intervals. The time between the various inspections depends on several factors. Certain measurements must be made at regular intervals because they are required by either the FCC Rules or the station license. An engineer taking over a station should immediately check the station license and construction permit to see just what measurements and measurement limits are set forth. If the monitoring-point measurements must be made weekly and the current ratios and phase angles must be held to tight tolerances, it is a sure bet that the Commission thinks that the array is critical—and they are usually right. Naturally, inspections must be made much more frequently with a critical array than with one that has loose tolerances.

Aside from the legal requirements, measurements and inspections must be made often enough to detect impending problems before they occur. How often is usually determined by experience. Certain inspections should always be made after a severe storm or whenever any of the regularly monitored parameters changes signifcantly.

MECHANICAL INSPECTION

A mechanical inspection is one that can be done without making electrical measurements. It is amazing how many troubles can be discovered by merely looking carefully at various parts of the system.

Starting with the towers themselves, such things as paint and lighting accessories should be checked regularly. The lightning gap and base insulator should be inspected regularly for changes. The separation of the gap may change. Or dirt and dust may collect on the base insulator and, after becoming wet, actually burn, causing a carbon track across the insulator. It is good practice to protect both the lightning gap and base insulator from the elements. The insulator can be given a thin coat of silicone to avoid the collection of moisture and a small rain shield can be installed to protect the gap.

The ground connection to the bottom of the base insulator and to the radials should be checked regularly. At least once a year the bolts on the tower should be checked. Another check to be made at least annually, and after every severe windstorm, is to use a surveyor's transit to assure that the towers are still vertical. The guy wires should be checked for tightness. Where cattle roam freely among the towers of an antenna, the guy wires are often used by the cattle to scratch their ever-present itches, which can cause turnbuckles to loosen. Protective sleeves can be put over the turnbuckles to avoid this problem.

The sampling loops should be inspected closeup or with field glasses. It helps if snapshots of the towers are available for comparison with the observations. Sampling lines, chokes, and tower-lighting lines should be inspected to make sure that they haven't been loosened by the wind or damaged by lightning. The area in the immediate vicinity of each tower should be free of vegetation, and the vegetation in the rest of the area should be kept mowed.

An item that is often neglected in inspection is the fence around the base of the tower. If the fence is metal, it should be firmly grounded at regular intervals. These grounds often work loose with time and can cause instability.

The next point for inspection, after the tower, is the "doghouse" at the base of the tower. All bolted connections should be checked to see that they are secure and free of any corrosion. The housing for the metering and impedance-matching networks must be kept clean. Mice, rats, and snakes seem to favor doghouses and other equipment housings for shelter. Often they disturb the setting of impedance-matching elements. Housings should be tight enough to avoid this situation, and if necessary, rat poison should be placed in the doghouse.

The feeder cables should be inspected regularly. If pressurized lines are used, the pressure should be checked and the gauges tapped lightly to make sure that they aren't stuck. If there is a leak in a line. it should be located and repaired as soon as possible, before moisture has a chance to penetrate the line. The ends of cables should be sealed against moisture.

If above-ground cables are used, they should be checked to see that the ground connections haven't come loose. There isn't much that can be done by way of visual inspection of underground cables except to inspect the ends and pressure. If trouble is indicated, it is usually necessary to dig up the cable.

The last item to be inspected visually is the phasor and power divider. Here again, all connections should be checked carefully. Taps on coils should be checked to see that they haven't come loose. It is helpful to mark the point on the coil where a tap is connected with fingernail polish so that if a tap should become loose and slip, it can be put back in its original position.

A harbinger of trouble in networks is often unusual heating of a coil or capacitor. A very useful tool for locating hot spots is a special crayon that changes color permanently at a predetermined temperature.* These crayons are available for different temperatures, and it is a good idea to put stripes around the capacitors and at various points on the coils, particularly near taps. Often a tap will loosen, causing a high resistance. with the temperature-sensitive marks, hot spots can be detected easily. Many capacitors of the types used in matching and impedance-transforming networks are actually made up of many smaller capacitors inside a housing. One or more of these small capacitors may deteriorate due to age or lightning surges. This type of trouble is almost always accompanied by a hot spot at some part of the case.

If coils or capacitors run hot all of the time, even when the array is operating properly, it indicates excessive loss, and the situation should be corrected promptly.

In summary, the mechanical inspection should include a careful look at all parts of the system to see if anything has changed since the last inspection. In spite of protective gaps. lightning surges do get into directional-antenna feeder systems. Sometimes the effects are obvious and can be recognized immediately. Other times the effects are more subtle, and it takes careful detective work to find them. In general, anything unusual should be noted in the maintenance record. A symptom might not seem important at the time it is observed, but it might provide a useful clue to the cause of trouble that develops later.

ELECTRICAL MEASUREMENTS

The FCC Rules require that the current ratios and phases, and usually the field intensity at monitoring points, be measured at regular intervals and the indications recorded. These measurements should be carefully reviewed as a part of the regular inspection procedure. It is important to keep records of the weather at the time the measurements are

*A wide variety of temperature-sensitive paints, labels, and crayons are available from Tempil Division of Big Three Industries, Inc., 2001 Hamilton Blvd., South Plainfield, NJ 07080.

made. If some measured parameter changes from its usual value to another value during a storm and doesn't return to its usual value after the storm is over, you can be pretty sure that something has been damaged by the storm. If, on the other hand, a measured parameter changes value during a storm and then gradually returns to its usual value, that is an indication that something gets wet during a storm and dries out when the storm is over. The culprit might be the base insulator of a tower, a transmission line, a component in the doghouse, or any of many other things.

DIRECTIONAL-ANTENNA STABILITY

The most frequent complaint of the broadcast engineer against directional-antenna systems is that they are unstable. In some cases, the design is critical, and there will always be some instability. However, in the majority of cases, an unstable array can be tamed considerably with a little analysis of the trouble spots and some corrective modifications. The difference between a stable antenna system and an unstable one is often merely a matter of the time and money spent on installation and maintenance. Usually the cost of corrective maintenance will be repaid in a short time by the time saved in troubleshooting and by improved signal coverage. In the next several pages we will look at the most common causes of instability and see some ways in which it may be corrected.

Variation in Ground Resistance

One cause of instability in an array is changes in the *loss resistance* at the towers. This resistance is almost all in the ground system. The effect of a variation in ground resistance is most pronounced in close-spaced arrays of short towers. The reason is that these arrays usually have low driving-point impedances at the bases of their towers, and the ground resistance is a substantial fraction of the total resistance.

One thing that can be done to minimize the effect of variations in ground resistance is to use tall towers in a wide-spaced array. This is of little interest to the engineer who has inherited a close-spaced array that uses short towers, but it is worth considering if it ever becomes necessary to move the station.

Once a system has been installed, there are four things that can be done to reduce the variation of ground resistance:

1. For each tower use at least 120 radials that are an average of 0.4 wavelength long. The radials for each tower should be connected as described in Chapter 12.
2. Install a square of copper mesh at least 24 by 24 feet at the base of each tower. All radials should be brazed to the screen.
3. Be sure that all metal fences and other conductors are kept well away from the base of each tower. If a metal fence is used, it should be bonded to the ground system at frequent intervals.
4. Keep all vegetation out of the area immediately adjacent to the tower. Vegetation over the rest of the ground system should be kept well mowed.

These steps will very frequently tame a system that has been notoriously unstable since its installation.

When an array that has been reasonably stable for some time after installation becomes unstable, the first thing to do is to make sure that there is still a ground system. In some types of soil the radials of old stations deteriorate to the extent that it is impossible to find any trace of them. Any part of the area covering the ground system that is heavily traveled should be checked to make sure that the radials haven't been broken. A sudden change in stability has in some cases been the result of one or more radials having been stolen. The presence of radials can be checked roughly with a field-intensity meter. Any sudden change in indication as the meter is carried over a radial may indicate that the radial has been broken at that point.

Many arrays have one or more towers that have a very low positive or negative driving-point impedance. These may sometimes swing back and forth between positive and negative. This means that the tower is taking very little power from the feeder system and is acting rather like a parasitic element. A small change in the ground resistance associated with such a tower will produce a rather large effect. It is sometimes helpful to install a fixed resistor in the matching network at the base of such a tower. This will tend to swamp the effect of changes in ground resistance. It will also increase the loss at the tower, usually such a tower handles a very small part of the total power, so the additional loss is not serious.

Temperature Effect

Changes in ambient temperature produce small changes in the effective length of transmission lines and hence in the phase shift that they introduce. This can be tolerated in some arrays, but in a critical array, every change is serious. In critical arrays it usually helps to bury all transmission lines so that they will not be exposed to changes in weather.

Most thermocouple ammeters are temperature-sensitive. When calibrated at room temperature, they may be in error by as much as 5% in extremes of temperature. When all of the base-current meters of an array are at the same temperature, this error does not have a serious effect on current ratios, but if for some reason one doghouse has a different temperature than the others, the meter error alone may be sufficient to indicate that the licensed tolerance has been exceeded. This situation is apt to occur when the transmitter building is located adjacent to one of the towers. The meter and impedance-matching network for that tower may be in the transmitter building, where the temperature is held reasonably constant; but the meters and matching networks for the other towers are in doghouses, which are not heated in the winter or cooled in the summer. In a critical array the components at the bases of all of the towers should be subject to the same temperature variations

When the network feeding one of the towers is in the transmitter building, the lead to the tower may be several feet longer than the leads to the other towers. As a rule of thumb, we can assume that a length of just about any wire has an inductance of about 1 μH per meter. This means that if the lead to one tower (after the matching network) is 15 ft longer than the leads to the other towers, a substantial amount of inductance will be added in series with the base impedance of that tower. A 15 ft length of wire has an inductive reactance of about 19 ohms at 600 kHz and 50 ohms at 1600 kHz. In a critical array this can cause problems. The problems can be minimized by adding an equivalent length of line at the other towers. The excess wire can be made functional by coiling it up into a small lightning-surge choke.

Networks and Instability

Networks are frequent contributors to instability. If the *Q* of a network is greater than about 3, the circulating currents will be excessive and the bandwidth of the network will be small. Large circulating currents heat components and change their value. This, in turn, changes the operating parameters of the array. If the required impedance-transformation ratio is greater than about 10, it is better to use two networks in tandem than to use a single network. In this way the *Q* can be kept low.

In the original design of a network where a small value of capacitive reactance is required, the arrangement of Fig. 16-3 is often used. Here a small capacitive reactance is obtained by using a small capacitor (high capacitive reactance) in series with an inductance. This arrangement is very economical because small capacitors cost less than large capacitors. If the principle is carried too far, however, the capacitive reactance will vary over a wide range when either of the components changed slightly. For example, in Fig. 16-3 we get 20 ohms of capacitive reactance by using 1000 ohms of capacitive reactance in series with 980 ohms of inductive reactance. At 1 MHz this means we can use a 0.0002 μF capacitor to get 50 ohms of capacitive reactance, rather than using 0.003 μF. At broadcast-antenna powers the saving is substantial. The arrangement seems, on the surface, to be ideal. However, the net reactance in the circuit of Fig. 16-3 is the difference between two large numbers, and such a quantity varies considerably when either of the large numbers varies only

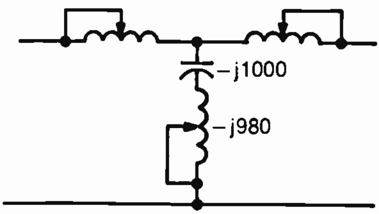


Fig. 16-3. Unstable shunt reactance.

slightly. For example, suppose that the reactance of the capacitor in Fig. 16-3 changed by 1% to 1010 ohms. The net capacitive reactance in the circuit would then change to 30 ohms, a change of 33%.

The chief reason for using the arrangement of Fig. 16-3 is that it is usually easier and more economical to make a variable inductor than a variable capacitor at the powers used in broadcasting. Another reason is that when the arrangement is used in the shunt arm of a *T*-network, it may be made series resonant at the second harmonic of the operating frequency, thus improving harmonic rejection. When the circuit is used for these reasons, the reactance of the capacitor should not be more than two or three times the desired value of net reactance of the circuit, if we are to avoid the problem mentioned in the preceding paragraph.

In checking a system for causes of instability, consider every place where inductors and capacitors are used in series as a possible culprit.

In practice, one of the most frequent causes of apparent instability is the antenna-monitoring system. All new antenna monitors are type-approved by the FCC, and they are usually very stable and accurate unless they have been damaged by lightning surges. Even the older antenna monitors, called *phase monitors*, are quite accurate. On the other hand, many sampling systems are anything but stable. When a sampling system is not stable, the engineer often thinks that his antenna array is not stable. This has caused many an engineer to retune an array that had nothing wrong with it and thus to put the monitoring-point parameters out of limits.

Whenever a current-ratio or phase indication is unstable, the engineer should make every effort to assure that the change is actually due to a change in the antenna instead of the monitoring system. If current ratios and phase angles change but the field-intensity measurements at the monitoring points do not, the sampling-and-monitoring system should be carefully checked.

When a monitoring system is found to be unstable, it should be repaired or completely replaced immediately. It will cause nothing but trouble, and the time and money required to replace it will be well spent.

Defective Transmission Lines

Sometimes defective transmission lines are at the root of instability. After everything else has been checked, and replaced as is necessary, the transmission lines should be checked. The following checks should be made on each line:

1. The open-circuit resistance
2. The short-circuit resistance
3. The line impedance

The first two checks are DC measurements made to locate improper resistors in the line or connectors. The impedance is measured to locate defects that will only be evident at high frequencies. All of these measurements should be recorded with a description of the measurement setup so that the measurements can be duplicated later to see if anything has changed.

Occasionally, a section of line is defective when received from the manufacturer. For example, the center conductor of a coaxial line may not be properly centered. In such a case, the high capacitance at the place where the line is defective causes a high current and, consequently, a hot spot. The heating may cause the characteristics of the line to vary, thus causing instability. Such a defect would cause an array to be unstable from the time it was installed.

In summary, some critical arrays exhibit a great deal of instability. Usually, however, an unstable array can be tamed considerably by making a very careful check for correctable causes of instability. This is definitely true if an array was stable when it was first adjusted.

RETUNING A DIRECTIONAL ANTENNA

Even with the most rigorous maintenance schedule and the best preventive maintenance, the dreaded day comes when the array must be adjusted to make it operate within the licensed parameters. Most broadcast engineers view tuning a directional antenna about the same way they view a game of Russian roulette. They feel that you may not lose, but you can't really win. This definitely isn't a job for a beginner; there are so many interacting variables in a directional-antenna system that many situations can be recognized only after many years of experience. On the other hand, if the job is approached from a sound engineering basis by an experienced engineer, the proper adjustments can be accomplished. In fact, once an engineer becomes familiar with the idiosyncrocies of a particular antenna system, he is often more proficient at adjusting it than even a consulting engineer.

Preliminary Steps

There are certain things that should be done before the adjustment of an array is attempted:

1. The engineer should make sure that the situation he is trying to correct is not caused by a component failure. If one of the components in the feeder system goes bad, perhaps some improvement can be attained by adjustment, but it is a losing battle. Sooner or later the component must be replaced, and then it will be necessary to put the adjustments back where they were. It is better to be sure that all of the components are in good condition before the adjustments are attempted.
2. The engineer should be familiar with the effect each control has on the pattern, also the amount of interaction he can expect between adjustments. This knowledge can be obtained by observing the consultant when he is tuning up or readjusting the array. There are other ways of becoming familiar with the action of the various controls, and these will be discussed later.
3. Before any control is touched, a careful record should be made of all dial indications, tap positions, etc. This will come in very handy if the adjustment procedure gets out of hand and it becomes necessary to restore the original setting.
4. It should always be remembered that the object of the procedure is to keep the pattern within its licensed limits. If possible, this should be done with current ratios and phase angles within the licensed tolerances, otherwise the FCC may require a complete proof of performance.
5. Whenever possible, the situation that indicates readjustment is required should be carefully analyzed

to find possible reasons for the situation. Only then can the adjustments be made intelligently.

There are several things that can be done to make the adjustment procedure much easier. The first is to determine just how each adjustment affects each current ratio and phase angle. This information can be gathered by carefully marking down the indication of a dial and then varying it in each direction while observing the effect on each current ratio and phase angle. The effect that each control has on the field intensity at each monitoring point can be obtained by an observer with a field-intensity meter at the monitoring point. As each control is varied on either side of its normal position, the observer can report the effect on field intensity. By carrying out these steps for each control on the power divider and phase shifter, a table can be made up that will show the effect of each control.

Suppose that an analysis as described above indicated that two controls had more influence on the field intensity at a particular monitoring point then any of the other controls. Suppose further that the field intensity at this monitoring point was out of limits. Then, logically, we would start the adjustment procedure by using the two controls that had the most effect. Probably the entire readjustment could be carried out with only these two controls.

Whenever any control of an array is varied, it might result in a change of the impedance at the common point. If it does, the common-point impedance should be adjusted before any further adjustments are made. This situation can be detected quickly if a common-point bridge is installed. Otherwise, the indications of the meters in the final stage of the transmitter should be checked. If the common-point impedance is disturbed very much, that will probably be shown by a change in the loading of the transmitter.

Usually, when an array requires readjustment, the required change in the pattern is small. If a major readjustment of several controls seems to be required, there is probably some other trouble that should be corrected before the adjustment is made. When the pattern is only slightly out of limits, the various controls should be carefully nudged to try to bring the pattern back within limits. Rarely can this be done by adjusting one control. In almost every array there is enough interaction that changing one control requires changing another.

TROUBLESHOOTING A DIRECTIONAL ANTENNA

If a careful maintenance program is in effect at a station, the probability of a catastrophic failure in the system is reduced considerably. Most components fail slowly. If they are watched on a regular basis, the impending failure can be detected and the component replaced beforehand.

Storm Damage

The most frequent cause of a catastrophic failure of any of the components of a directional-antenna is a thunderstorm or a severe windstorm. Damage of this type is often found by visually inspecting the system. Suppose, for example, that a system has been struck by lightning and has failed. The place to start looking for the trouble would be at the towers. Lightning enters the system through the towers and usually does the most damage there. The lightning tries to establish a path between a cloud and the ground. Unfortunately, some of the charge often gets back as far as the phasor, and it may destroy any component that stands in its way. If all of the components in the doghouse check out okay, each of the lines should be checked. Finally, the components in the phasor should be checked.

Windstorms usually affect components on the tower. Sampling loops may be loosened or moved, or lines may be moved and connections broken loose. Here again, any damage is usually visible.

A more insidious effect of lightning is partial damage to a component. A capacitor in a matching network may become partially shorted. The effect noticed at the time may be slight, but conditions will only grow worse as time passes. If the regular maintenance program requires following up any small changes in meter indications to find what component of the system has changed, this type of trouble will be located quickly.

Changes in Antenna-Monitor Readings

Usually, the first sign of trouble in a directional-antenna system is a change in tower currents or phases. The first thing to do when current ratios or phases go out of limits is to check the field intensity at the monitoring points to see if the pattern has changed. If it hasn't, the trouble is likely to be in the monitoring system rather than in the antenna itself. If the pattern has changed, the way in which it has changed may give a clue as to the trouble.

Suppose for the moment that the trouble is most noticeable in the ratio and phase in a particular tower. Because of the interaction between towers, there will almost always be some indication of trouble in the feeder lines of other towers. But if a component is faulty in one of the feeder lines, the indications pertaining to its tower will probably be out more than the other, tower indications. Suppose further that a change in monitoring-point measurements also indicates that the trouble is in the suspected tower. The next step is a careful visual inspection, not only of the tower itself but of all the components in its feeder system. Such things as faulty connections. loosened nuts and bolts. and loose or shifted coil taps can be spotted, and the situation can then be corrected.

If the visual inspection doesn't show anything wrong, electrical tests must be made. Here is where a complete diagram of the system showing the measured values of all of the components will be worth its weight in gold. The impedances can be checked. starting at the input to the base-matching network of the tower. If this impedance is okay, the measurements may be made back at the phasor. If at any point the impedance is far from its normal value, the values of the individual components may be checked with a bridge. This will usually disclose the source of the trouble.

When the trouble seems traceable to one tower and yet all of the line impedances and component values are normal, look for something wrong in the grounding of that particular tower.

Occasionally a situation is encountered where, without warning, all of the current ratios and phase angles in the entire array change drastically. At first glance it appears that the entire array has suddenly become defective. One common cause for such a situation is a defect in the antenna monitor that renders all of the indications incorrect. The first check for this is to take actual base-current measurements at each of the towers. If the ratios as calculated from the base-current measurements are okay, the trouble probably is in the monitor. The manufacturer's instructions should be followed for troubleshooting the monitor.

Another very common cause for all ratio and phase readings being wrong is trouble in some component in the line feeding the reference tower of the array. All of the current ratios and phase angles displayed on the antenna monitor are related to the magnitude and phase of the current feeding the reference tower. If the magnitude and phase of this current are not correct, all of the indications of the antenna monitor will be incorrect. Some troubles of this type can be recognized by temporarily connecting the antenna monitor to use one of the other towers as a reference. For example, if the current ratios of towers 2 and 3 with reference to tower 1 are both 0.6 then, when tower 2 is used as the reference, the ratio at tower 3 should be 1. This check won't always work, but it is worth a try.

Whenever an array is adjusted to bring the current ratios and phase angles back to their licensed values, the field intensity at each of the monitoring points should be checked. Unfortunately, in many arrays, the ratios and phases can be brought back to their licensed values, but the pattern will still be out of limits.

Monitoring Points Out of Limits

Another signal for antenna readjustment is that the field intensity at one or more of the monitoring points is out of limits. When the field intensity is higher than the MEOV for a particular monitoring point, the situation is most serious. If an inspector from the FCC was to arrive on the scene, he would issue a citation. Whenever the indications at any of the monitoring points exceed their maximum licensed values, a good first step is to call the FCC. Explain what has been observed and what steps are being taken to find and correct the trouble. The Commission is usually cooperative in these matters, and usually some way can be found to keep the station on the air legally while the trouble is being investigated. Depending on how critical the situation is, the station may receive telegraphic authority to operate with station parameters at variance with their licensed values, or possibly permission to operate with reduced power.

When the field intensity at a monitoring point has dropped, the situation is not as serious, in that it won't cause interference to the service area of another station, but the signal may not be satisfactory in the station's own service area. In any case, the trouble should be found and corrected as quickly as possible.

In northern latitudes a frequent cause of incorrect monitoring-point indications is the change in ground conductivity due to weather conditions. This trouble rarely appears suddenly, but is noticed as a gradual change in the field intensity, usually at all of the monitoring points. The array should not be readjusted; the trouble should be resolved some other way.

Rarely will the field intensity decrease at all of the monitoring points unless something is wrong before the common point. The more usual situation is that the field intensity decreases at one or two of the monitoring points and increases at one or more of the others. Before doing anything other than performing a visual inspection, it is a good idea to analyze the situation to see how the current magnitude and phase in each tower affect the field intensity at each monitoring point. In fact, it is a good idea to have such an analysis available anyway, for it will save a great deal of time when it is needed. The best way to analyze the contribution of each tower to the field intensity at each monitoring point is to make a separate vector diagram of field intensity for each monitoring point.

Chapter 17

FM and TV Antennas

There are many basic differences between the antennas that are used in the standard broadcast band and those that are used for FM and TV broadcasting. From the point of view of the engineer who is charged with operating and maintaining them, the main differences stem from the frequencies that are used. FM and TV broadcasting is done at VHF and UHF, where the dimensions of an antenna are small enough that a complete antenna can be fabricated by a manufacturer and shipped to a station. In contrast, standard broadcast antennas are so large that the antenna and its ground system may occupy more than a square mile of space. Thus the standard broadcast antenna is made up of components that are often made by different manufacturers and assembled on the spot where the antenna is to be used. The job of the broadcast engineer is to keep all of the components working together as they were intended to. For an FM or TV station the entire antenna-and-feeder system may be made by the same manufacturer and shipped to the station nearly intact. Its maintenance is a whole new ball game for the engineer.

Other differences stem from the service area involved. FM and TV service areas are protected simply by geographic separation. The signals normally travel in a line of sight, and if the stations are far enough apart, they won't interfere with each other. In the standard broadcast band service-area protection is accomplished by appropriate design of the shape and size of antenna patterns.

The FM broadcast band occupies frequencies between 88 and 108 MHz. It is divided into 100 channels, each of which is 200 kHz wide, thus FM broadcasting is a narrowband service. TV broadcasting occupies 82 channels, each of which is 6 MHz wide. The TV channels lie in the four bands 54–72, 76–88, 174–216, and 470–890 MHz. TV is a wideband service; in fact, the requirement for wider bandwidth is one of the major differences between TV and FM antennas. It is important that the impedance at the terminals of the antenna remain reasonably constant over at least 10% of the bandwidth for TV and 0.2% of the bandwidth for FM. Some FM antennas are designed with a wider bandwidth so that the same antenna may be for used for more than one FM channel. This reduces the number of types of antennas that a manufacturer has to produce.

In both the FM and TV services the usual practice is to use a horizontal and essentially circular radiation pattern. The FCC does not allow directional patterns in the horizontal plane for the purpose of reducing the required separation between stations on the same channel, but it does allow some directional gain to provide better coverage of populous areas. The signal in one direction may exceed that in another by no more than 10 dB.

DEFINITIONS

There are several terms used in connection with FM and TV antennas that haven't been used earlier in this book. Some of them are defined here.

Antenna Height Above Average Terrain. This term is a measure of the height of the antenna with respect to the surrounding terrain. To find its values the average height of the terrain at distances of 2 to 10 miles from the antenna is found; then the height of the antenna above this average is taken. This procedure is repeated along eight radials, for each 45° of azimuth, starting with true north. In general, a different average antenna height will be found along each radial. The average of these heights is the antenna height above average terrain. Sometimes, fewer than eight radials are used when some of the area is over large bodies of water or outside of the United States.

Where the antenna uses circular or elliptical polarization, the height is measured from the center of radiation of the horizontal component of the field.

Antenna Power Gain. In FM and TV broadcasting the antenna gain is specified with respect to the gain of a half-wave dipole. The pattern of a horizontally oriented half-wave dipole is such that for 1 kW of radiated power the field intensity on the surface of the earth is 137.6 mV/m. Therefore, the gain of an antenna in decibels is given by the formula

$$\text{dB} = 20 \log (E/137.6)$$

where field intensity is mV/m.

Decibels Referred to One Kilowatt, dBk. This is a logarithmic measure of power level. It is the number of decibels by which a signal is greater than 1 kW and is given by

$$n_{dBk} = 10 \log P$$

where P is power in kilowatts.

Decibels Referred to One Microvolt Per Meter, dBμ. This logarithmic measure field intensity is the number of decibels by which given field intensity is greater than 1 μV/m. It is given by

$$n_{dBu} = 20 \log E$$

where E is the field intensity in microvolts per meter.

Effective Radiated Power, ERP. The effective radiated power is the product of the antenna power and the antenna gain. The antenna power as used here is equal to the transmitter power less the losses in the transmission lines leading to the antenna. It is frequently expressed in decibels referred to 1 kW (dBk). When a circularly polarized antenna is used, an effective radiated power is applied separately to the vertical and horizontal components of the radiation.

ANTENNA HEIGHT AND GAIN

FM and TV signals travel in a straight, line-of-sight path. For this reason, the area served depends on the transmitter, the height of the antenna, and the gain of the antenna. The horizontal-radiation pattern is usually circular or nearly so. Since protection of cochannel stations is accomplished by geographical separation, there is no need for directivity in the horizontal plane. In fact, the FCC Rules will not allow a directional pattern for the purpose of reducing the minimum separation between stations on the same or adjacent channels. Some directivity is allowed to permit the use of a desirable antenna site, but the ratio of maximum to minimum radiation in any direction can't be greater than 10 dB.

The *gain* of an FM or TV antenna is specified with respect to the radiation from a half-wave dipole. If an antenna has gain, that means that the antenna radiates more energy into some region than would be radiated into the same region by a half-wave dipole. Because of the law of conservation of energy, this means that the antenna radiates less energy into some other region than would be radiated by a half-wave dipole. We don't get something for nothing.

At the frequencies used for FM and TV broadcasting, half-wave elements are small enough that many of them can be mounted on a single tower. This permits us to stack antenna elements vertically and thus squeeze the pattern so that more energy is radiated in the horizontal plane and less at vertical angles.

The FCC Rules regulate the effective radiated power of FM and TV stations. Thus the desired coverage can be obtained with some combination of transmitter power and antenna gain. At the high UHF TV channels, transmitter power is expensive, so antenna gains as high as 60 are found at these frequencies.

The Rules also provide that a minimum field intensity be provided over the principal community which the station is licensed to serve. FM stations must provide a field intensity of 3000–5000 μV/m over the community. The minimum required field intensity from TV stations depends on the channels used as shown below.

Channels	*Minimum Field Intensity (dBμ)*
2– 6	74
7–13	77
14–83	80

When the antenna site is on a hill and much of the principal community being served is in a valley, providing the minimum required field intensity is as much of a problem as providing coverage of the outer limits of the service area.

RADIATING ELEMENTS

FM and TV antennas come in many different configurations, and it is often difficult to see any resemblance to any familiar antenna element. Actually, most FM and TV antennas use elements that are based on four fundamental radiating structures: the dipole, the loop, the slot, and the helix. We will review each of these to see how it works and how it can be used to produce the desired pattern and gain.

Figure 17-1 shows a half-wave dipole and its radiation pattern. We touched on this type of antenna element earlier and saw that it has a terminal impedance of 72 ohms at the frequency at which it is resonant. The dipole looks electrically like a series-resonant circuit. Thus its terminal impedance is resistive only at one frequency. Fortunately, if the diameter of the arms of the dipole is increased, the inductance is decreased and the capacitance is increased. This has the same effect as lowering the Q of a resonant circuit. It tends to reduce the rate of change of impedance with frequency and thus increases the bandwidth. It is for this reason that TV antenna elements are made thick.

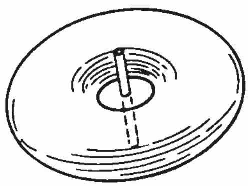


Fig. 17-1. Radiation pattern of a half-wave dipole.

Figure 17-2 shows some antenna elements that are similar in operation to the dipole. Figure 17-2A shows the biconical antenna, in which the arms flare out toward the ends. This configuration has a very wide bandwidth. Figure 17-2B shows a folded dipole. In this a second element is mounted parallel to the dipole, and the two are connected together at the ends. The current and voltage distributions are about the same in both elements. As a result the impedance at the terminals is about

300 ohms, versus 72 ohms for a regular dipole. The bandwidth of the folded dipole is greater than that of the regular dipole because the folded dipole has, in effect, a greater diameter. The input impedance can be increased still further by adding additional elements as shown in Fig. 17-2C.

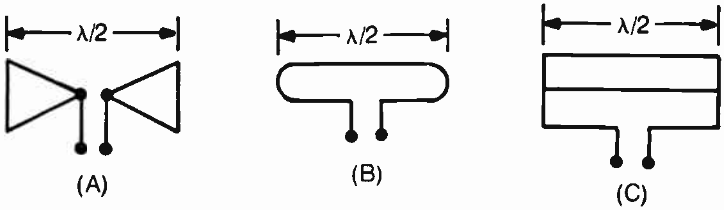


Fig. 17-2. Basic half-wave antenna elements.

Figure 17-3 shows a loop antenna. Though the ordinary half-wave dipole can be thought of as an *electric dipole*. because the ends have an opposite electric charge at any instant, it is convenient to think of the loop as a *magnetic dipole*. The loop may be either square or round. In fact, any loop of wire will radiate if the currents in opposite sides are not equal and opposite and very close together. Radiation is always zero in a direction perpendicular to the loop, but a wide variety of patterns can be provided by various configurations of loops.

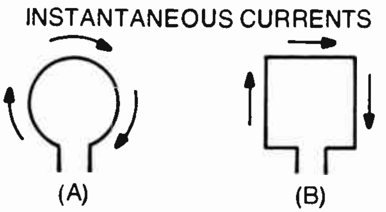


Fig. 17-3. Elementary single-turn loops.

The slot antenna is shown in Fig. 17-4A. Part B of the figure shows a dipole antenna that has been cut out of a sheet of metal. Of course, if we fed energy to it, it would radiate with the familiar dipole pattern. What isn't so obvious is that if we connected leads to the hole left in the sheet when we cut out the dipole as, in Fig. 17-4B, the hole (slot) would also radiate. With the sheet positioned as shown, it would radiate in the horizontal plane with the same pattern as the dipole. The slot antenna has about the same bandwidth as the dipole. The input impedance of the slot is related to that of the dipole by the formula

$$Z_{slot} = \frac{35{,}476}{Z_{dipole}}$$

Thus, since the impedance seen looking into the dipole is 72 ohms, the impedance seen across the center of the slot is 493 ohms. A great deal of variation in the input impedance of the slot can be had by changing the place where it is fed.

Strictly speaking, what has been said about the slot antenna is true only if the sheet from which the dipole is cut is considered infinite in extent. Current flows in the entire sheet, and since it is not infinite in extent, there will be reflections from the edges that interfere with the neat picture painted here. If the sheet is large compared to the slot, however, most of the energy will be radiated before it reaches the edges of the sheet so there won't be much reflection. Frequently the sheet is bent into a cylinder.

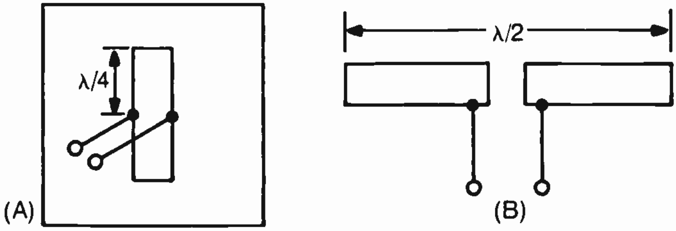


Fig. 17-4. Slot antenna.

Figure 17-5 shows a conductor wound into a helix. This element radiates either along the axis of the helix or at right angles to the axis. When the diameter and pitch are properly chosen, the radiation is at right angles to the axis of the helix. Helical antennas have been used extensively in TV broadcasting.

POLARIZATION

Traditionally, horizontal polarization has been used for both FM and TV broadcasting. In recent years circular polarization has been used increasingly in FM broadcasting to reducing the fading that is otherwise encountered in vehicular reception. As automobile FM receivers become more popular,

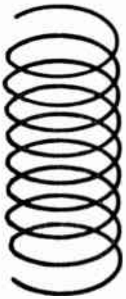


Fig. 17-5. The helix, basis of the helical TV antenna.

this use of circular polarization may be expected to increase. Experimental use of circular polarization in TV broadcasting shows great promise for reducing "ghosts" (spurious images in received pictures) in urban areas where there are many reflections. When a circularly polarized signal is reflected, its sense is reversed. That is, if a right-hand circularly polarized signal is transmitted, it becomes upon reflection from a building or other structure a left-hand circularly polarized signal. A receiving antenna of one polarization is "blind" to a signal having an opposite polarization. Thus, receiving antennas designed to receive right-hand-polarized signals will not see the left-hand reflected signals that cause ghosts. It is too early to tell whether or not this use of circularization in TV will become widespread.

GAIN REFERENCE

The gain of an FM or TV antenna is always referred to as a half-wave dipole. That is, the gain in a particular direction is a measure of how much more signal an antenna radiates in that direction than a half-dipole would. FM and TV allocations are based on horizontal polarization. The reference is thus a horizontal dipole. Figure 17-6 shows an end view of a horizontal

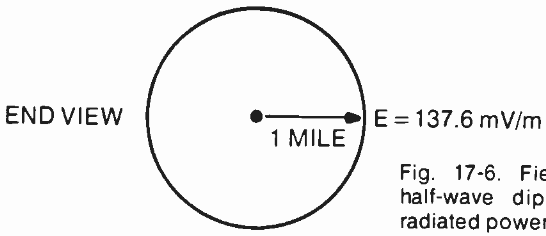


Fig. 17-6. Field intensity from half-wave dipole for 1 kW of radiated power.

half-wave dipole in free space. For the present, we will ignore the effect of reflections from the ground. The pattern has a field intensity in the horizontal plane at one mile from the antenna of 137.6 mV/m for a radiated power of 1 kW. The field intensity is proportional to the square root of the radiated power, so we can find the field intensity for higher powers by multiplying by the square root of the power in kilowatts. For example, if the radiated power was 50 kW, the radiated field intensity at one mile from the antenna would be

$$E = 137.6\sqrt{50} = 973 \text{ mV/m}$$

OMNIDIRECTIONAL PATTERN

In AM broadcasting we use a single vertical radiator to obtain an omnidirectional pattern in the horizontal plane and a field that is vertically polarized. In FM and TV broadcasting we want an omnidirectional pattern, but we want the field to be horizontally polarized. Thus we cannot use the simple expedient of a single vertical element.

Figure 17-7 shows a top view of a horizontal dipole and its radiation pattern. This pattern is far from omnidirectional. Now suppose that we were to use two dipoles at right angles to each other. Figure 17-8A pictures the fields that the two dipoles would produce if they were acting alone. Each dipole would have a pattern like that shown in Fig. 17-7. The way in which these patterns combine to form a resultant depends on the relative phase of the currents in each of the dipoles.

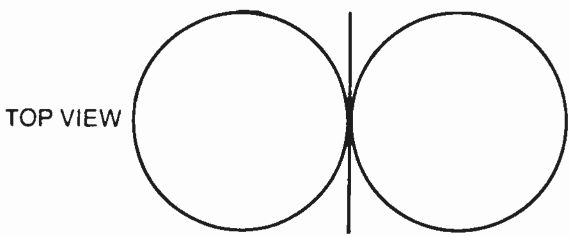


Fig. 17-7. Pattern of half-wave dipole (horizontal plane).

If the two dipoles of Fig. 17-8A were fed in phase, the fields would add to produce the figure-eight pattern shown in Fig. 17-8B. This isn't any better for our purposes than the field of a single dipole. But what if we excite the two dipoles 90° out of phase with each other? The fields in time quadrature and at right angles combine in somewhat the way oscilloscope signals in time quadrature and applied to deflection plates at right angles combine to produce a circular Lissajous pattern. The turnstile array thus produces a pattern that is very nearly circular in the plane of the turnstile.

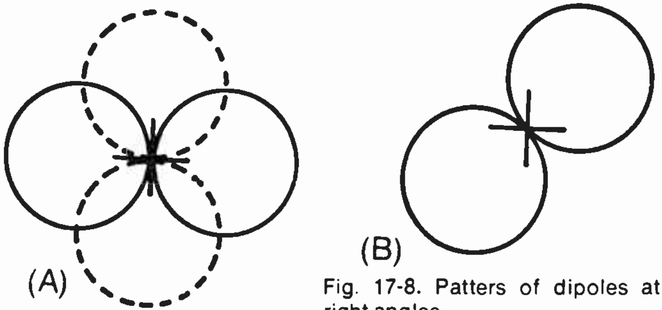


Fig. 17-8. Patters of dipoles at right angles.

If the turnstile were in free space, without any conducting structures nearby, and if the elements were very short, the pattern would be a perfect circle. In practice, the elements are 1/2 wavelength long and there is always a supporting structure nearby. As a result, the pattern isn't a perfect circle but is somewhat as shown in Fig. 17-9. This pattern deviates from true circularity, but it can be made good enough for most purposes.

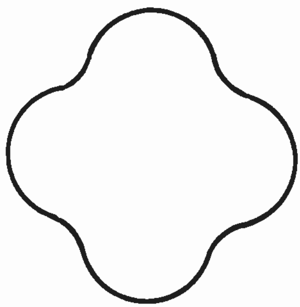


Fig. 17-9. Typical pattern of turnstile antenna.

Another method that is used to get an omnidirectional pattern is to fold a dipole into a circle. Figure 17-10A shows a folded dipole together with the normal current distribution. The current is maximum at the feed point and minimum at the ends. In Fig. 17-10B the folded dipole is bent into a circle. This antenna tends to radiate in all directions, but the radiation is not uniform, because the current is much greater on the side where the antenna is fed than on the opposite side, near the ends. The situation is improved in Fig. 17-10C, where the ends are brought close together and fitted with capacitor plates. The capacitive coupling tends to increase the current at the ends of the dipole. This makes the current distribution nearly uniform all around the antenna, and this in turn tends to make the pattern omnidirectional in the horizontal plane. By adjusting the capacitance at the ends of the dipole, the antenna can be tuned and the impedance adjusted.

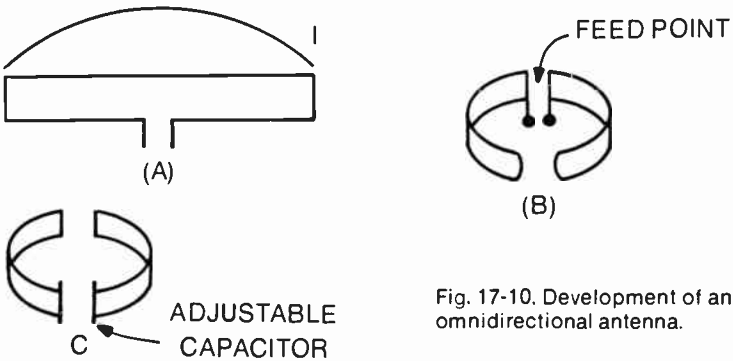


Fig. 17-10. Development of an omnidirectional antenna.

ANTENNA ARRAYS

In FM and TV antennas we use arrays of elements to control the directivity of the vertical-radiation pattern. Because of the comparatively small size of FM and TV antennas we can stack many of them on a single tower. Figure 17-11A shows the vertical-radiation pattern of a typical antenna. Note that since the signal propagates in a line-of-sight manner and is not reflected back to the earth, all of the energy radiated above the horizontal axis of the pattern is lost. Two steps can be taken to minimize this lost energy—stacking elements and tilting the beam. Figure 17-11B shows how the pattern is narrowed when several elements are stacked. Note that much more energy is transmitted along the horizontal axis, but that all above the horizontal axis is still lost.

Figure 17-12A shows the pattern of a half-wave dipole as viewed from the end. It is assumed that the dipole is far

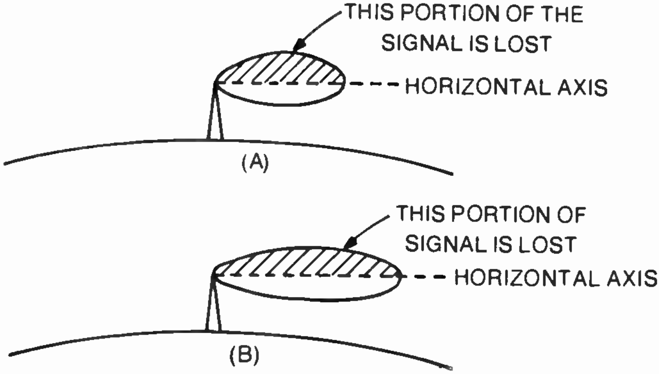


Fig. 17-11. Vertical-antenna patterns.

enough from the ground that there will be no reflection. The pattern is thus a circle. It is similar to the pattern that we would see if we were to look directly down on the top of a vertical antenna. In Fig. 17-12B we have another half-wave dipole one wavelength above the first one. The pattern is now the vector sum of the field intensities of the two dipoles. Note that the pattern has been flattened, with more energy being radiated horizontally and less being radiated at vertical angles. In fact, the radiation is zero directly above and below the antennas.

Figure 17-12C shows the pattern that results when six dipoles are stacked 1/4 wavelength apart. The pattern shows that most of the radiation is along the horizontal axis, which is desirable, but the pattern is still not optimum. For one thing, there are several nulls in the pattern at angles below the

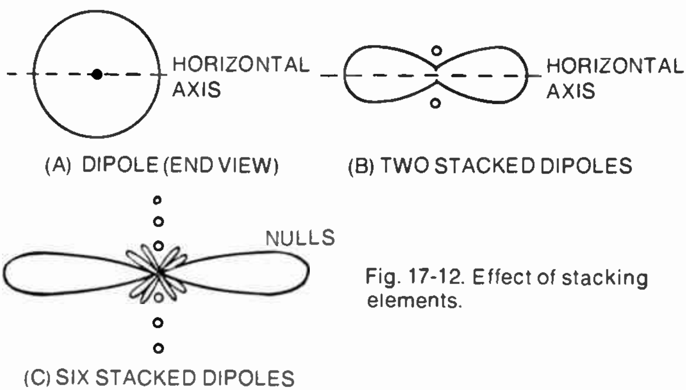


Fig. 17-12. Effect of stacking elements.

horizontal axis. Furthermore, half of the energy is radiated at angles above the horizontal axis, where it will never be received. We will first look at ways of removing the nulls from the pattern.

Null Filling

In our investigation of AM directional antennas we found that true nulls occur in a pattern because, at certain angles, the field from one antenna subtracts from that of another antenna so that the signal cancels completely. We found that nulls can be partially filled by making the currents in the various elements unequal. We have a very similar situation in the pattern of Fig. 17-12C. Here the nulls are not intentional, but result from stacking elements vertically to increase the radiation along the horizontal axis. The nulls exist because at certain angles the field intensity from one or more elements completely cancels that from the other elements. If the currents in the elements were not equal, their fields would not completely cancel, so the answer to the null problem is to feed unequal currents to the antenna elements.

Figure 17-13A shows a 6-element antenna with the elements second from the top and bottom producing only half of the field intensity produced by the other elements. There are no nulls in the pattern except directly above and below the antenna assembly. This is the technique of null filling by proper distribution of power. The pattern of Fig. 17-13A is the best so far, but half of the radiation is still lost at angles above the horizontal axis. This situation can be improved by beam tilting.

Beam Tilting

One way to accomplish a certain amount of beam tilting is to simply tilt the antenna itself slightly. Another way to accomplish beam tilting is to tilt the radiation from the antenna electrically. This is done by properly controlling the phase angle of the current in each element of the array. Figure 17-13B shows the same 6-element array, but now the phase of the currents is such that the fields add in a direction slightly toward the earth. In the figure the phase shifts are such as to tilt the beam 2° toward the earth.

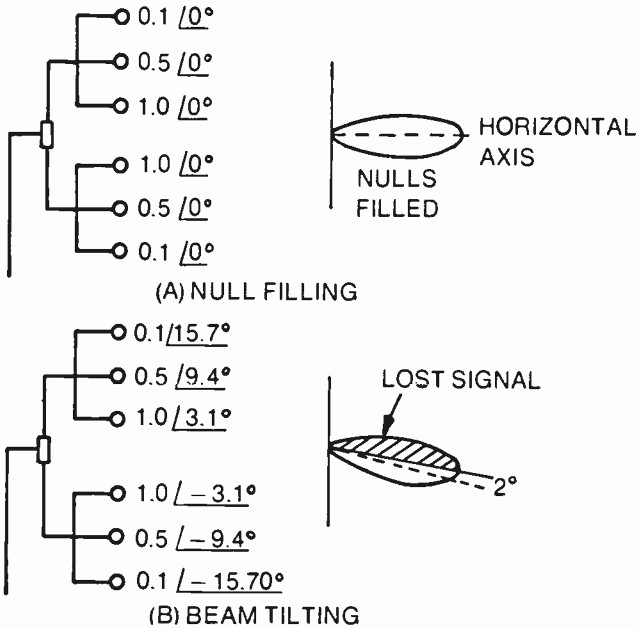


Fig. 17-13. Null filling and beam tilting.

Beam tilting is necessary not only to prevent radiation from being lost at vertical angles but also to make sure that the main lobe of the pattern actually reaches the surface of the earth. Because of the curvature of the earth, the maximum radiation of the antenna of Fig. 17-13A would never actually reach the earth. If an antenna is 1000 ft above the earth, the beam must be tilted at least 0.5° to reach the surface of the earth. Antennas on hills or mountains often use up to 2° of beam tilting for this reason.

SIDEMOUNTING

It once was assumed to be necessary to mount an FM or TV antenna on top of a tower, where there would be no supporting structure in the plane of the antenna to interfere with the radiation pattern. This is not always feasible, however, and at some stations both FM and TV antennas are mounted on a tower that is also used for AM broadcasting.

It has been found that remarkably good radiation and impedance characteristics can be obtained with various types of FM antennas mounted on the side of a tower. Undoubtedly, the antenna is detuned somewhat by side mounting, but the FM signal occupies a comparatively narrow band, and the antenna will often function very well.

ANTENNA-PATTERN PLOTS

With an AM antenna we use a polar plot of the pattern since we are concerned with how the signal distributes over the earth's surface. With an FM or TV antenna we are concerned with how and at what point the radiation from the antenna reaches the earth's surface. For this purpose it is often convenient to plot the pattern using rectangular coordinates. A rectangular plot of an antenna pattern is shown in Fig. 17-14. The vertical axis of the graph is graduated in terms of field intensity, and the horizontal axis is graduated in degrees below the horizontal axis through the antenna. This type of plot makes it easy to find the signal strength at any point along the surface of the earth. We know the height of the antenna above the ground, so we can calculate the angle at which the energy must leave the antenna to reach the surface of the earth at any given distance from the antenna. Many manufacturers specify the patterns of their antennas in rectangular plots.

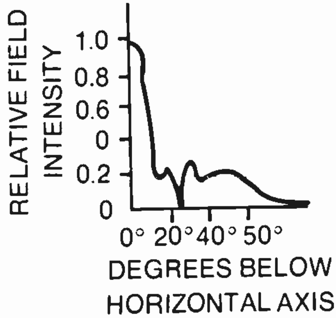


Fig. 17-14. Rectangular antenna pattern.

CIRCULAR POLARIZATION

Many FM stations are now using circular polarization, and some TV stations are experimenting with it. To produce circular polarization, an antenna usually radiates both horizontally and vertically polarized signals. To radiate vertically polarized signals with a dipole element, it must be oriented vertically. Thus many FM and TV antennas contain both vertical and horizontal elements. Figure 17-15 shows an example. Energy is fed to the elements by means of a power divider that is made of transmission-line elements. With this arrangement, only one transmission line from the transmitter is required.

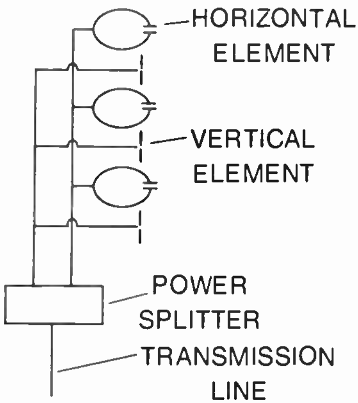


Fig. 17-15. Circularly polarized antenna with both horizontal and vertical elements.

PRACTICAL FM AND TV ANTENNAS

The FM or TV antenna, unlike a standard broadcast antenna, is furnished as a complete product by the manufacturer. In this respect, it is like an amplifier or transmitter. Usually, the broadcast engineer doesn't even have to concern himself with exactly how the antenna works. When he does become interested, all of the necessary information is in the manufacturer's literature. This is the opposite of the situation with an AM antenna for which there is usually no literature and no instruction book. The operation of any FM or TV antenna can be understood by extension of the principles of the basic antenna elements described earlier in this chapter.

The main problems that the engineer will face regarding an FM or TV antenna involve the feeder system and the associated measurements. These matters are covered in the following chapters.

Chapter 18
FM and TV Antenna Feeder Systems

GENERAL CONSIDERATIONS

As in any other type of antenna system, the principal objective of the feeder system of an FM or TV antenna is to provide an efficient transfer of energy from the transmitter to the antenna. Because of the high frequencies used, losses on transmission lines are higher and impedance matching over the necessary bandwidth is essential. This means that the standing-wave ratio on the transmission line must be kept low. FM stations transmitting only a single program with no subcarrier or stereo normally operate with a VSWR of 1.5:1 or less. FM stations carrying stereo or subcarrier broadcasts, as well as TV stations, operate with a VSWR of 1.1:1. The reason for the low VSWR is that serious reflections on the line would seriously distort the program material.

If the driving-point impedance of an antenna does not match the characteristic impedance of the transmission line, there is a reflection back along the line. There is usually a mismatch of some magnitude at the sending end of the line, so the reflected wave will be re-reflected back toward the antenna. This reflected signal thus bounces back and forth along the line until it is radiated or is dissipated in the line. The portion that is dissipated in the line increases the loss, but otherwise it does no harm. In the case of a TV station, the radiated portion of the reflected signal appears as one or more ghosts on the viewers' receivers. The magnitude of the ghost signal depends on the magnitude of the reflection, and the amount by which the ghost image is displaced from the original image depends on the length of the line. Naturally the situation becomes more objectionable when the line is long and the reflection strong. TV antennas are usually mounted on tall towers to provide good coverage, so reflections are always troublesome from this point of view. FM stations carrying stereo or subcarrier broadcasts experience reflections that manifest themselves as crosstalk between the channels.

FM and TV stations always use coaxial cables as transmission lines, except for some UHF stations, which use waveguide. The principles of the feeder system are the same in either case.

The choice of a particular transmission line depends on the transmitted power, the frequency of operation, and the length of the line. The line must have a power-handling capability at least as great as the total output power of the transmitter. The line should be rated for use at the operating frequency. Whenever the line length is more than a few hundred feet, the power dissipated in the line must be taken into consideration.

Coaxial cables used for FM and TV transmission lines usually have a characteristic impedance of either 50 or 75 ohms. They are always pressurized when run out of doors. Rigid lines are normally supplied in lengths of 20 ft. The lengths are joined by the use of flanges. No matter how carefully the sections are assembled, there will always be a slight impedance mismatch at the joint. At certain frequencies the reflections from the ends of the 20 ft sections will add to cause a serious standing wave on the line. Lengths must be selected to prevent cumulative reflections.

Vertical runs of line expand because of heating of the line. This expansion is allowed for by mounting the top of the line securely and supporting it on spring hangers that are attached to the tower.

The impedance of an FM or TV antenna may vary over a wide range, depending on the particular type of antenna. It is extremely important to match the impedance to the characteristic impedance of the line. At FM and TV frequencies it is difficult to use the same types of networks that are used in the standard broadcast band because the required values of inductance and capacitance are very small. For this reason, transmission-line sections called *stubs* are usually used for impedance matching.

IMPEDANCE-MATCHING STUBS

An open or shorted section of transmission line has a driving-point impedance that looks like a pure reactance or susceptance. By selecting the characteristic impedance of a section of line, as well as its length, the engineer can control both the value of reactance seen at its terminals and the rate, or slope, at which the reactance varies with frequency. The slope of reactance versus frequency is always positive and is greater than that of a lumped inductance or capacitance.

Instead of controlling the value and slope of reactance, the engineer can instead control the value of reactance at two frequencies.

We saw earlier that we can transform any value of impedance into any other value by using an *L*-network consisting of a single inductance and a single capacitance. We can do the same thing with a stub connected across a transmission line . In Fig. 18-1 the reactance of section 1 in part A corresponds to the series reactance of the *L*-network in B, and the reactance of the stub section, corresponds to the shunt reactance of the *L*-network.

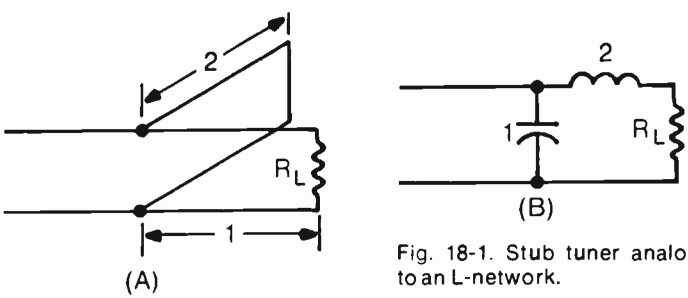


Fig. 18-1. Stub tuner analogous to an L-network.

In theory either a shorted or open section of line may be used as a stub, but in practice a shorted section is always used. There are two reasons. The shorted stub doesn't radiate anywhere near as much as an open stub, and the length of a shorted stub can be varied easily by means of a shorting bar or disc. In considering that a shorted stub had a driving-point impedance that looked like a pure reactance, we neglected the losses in the stub. There are always losses, but these can nearly always be neglected. An impedance transformation accomplished by transmission-line sections is affected much less by the unavoidable losses than that obtained by a conventional network.

There are many formulas that can be used to find the lengths of line used in matching-stub arrangements, but the simplest approach is to use the Smith chart to solve the problem. To solve stub problems, we must use admittance rather than impedance because the stub, by its very nature, is connected in parallel with the line. This is no problem, because, if the nature of the load is specified in terms of its impedance, we can transform this into admittance right on the Smith chart. Another advantage of using the Smith chart is that we don't even have to know the value of the load impedance or admittance. All we have to know is the standing-wave ratio and either the distance along the line to the load or the distance to a point on the line where the voltage is maximum or minimum.

Single-Stub Problem

Assume that we have a load impedance of $100 + j100$ ohms and wish to use a single stub to match it to a transmission line having a characteristic impedance of 50 ohms. The first step is to normalize the load impedance to 50 ohms.

$$\frac{100}{50} + \frac{j100}{50} = 2 + j2$$

Next, by using a lengthy formula and a calculator, or a Smith chart, we find that the normalized admittance of the load is $0.25 - j0.25$. Of course, we want the driving point of our matching arrangement to have a normalized admittance of $1 + j0$. In Fig. 18-2 we have the point $0.25 - j0.25$ plotted on a Smith chart. We also have the locus of points corresponding to a conductance of 1, or the unit-conductance circle. Any admittance that lies on this circle can be converted to an admittance of 1 + $j0$ by the susceptance of our shorted stub. The admittance seen across the line varies with distance from the load by taking on the values on a circle of constant standing-wave ratio. This is the dashed circle in Fig. 18-2. From this figure we can see that the admittance across the line intercepts the unit-conductance circle at a distance of 0.219 wavelength from the load. The admittance at this point is 1 + $j1.6$. If we continue along the dashed circle of constant standing-wave ratio, we find that it intercepts the unit-conductance circle against 0.362 wavelength from the load. This means that we could connect our stub, which has a pure susceptance, at either of these points to match the load to the line. Of course, there are other points like this each 1/2 wavelength farther along the line from the load, as the impedance of a transmission line (neglecting losses) repeats

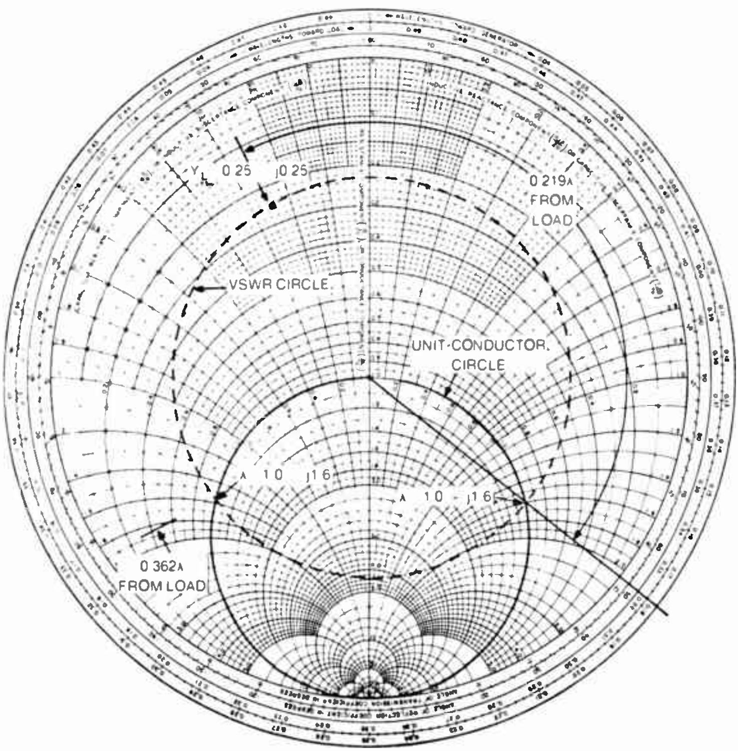


Fig. 18-2. Solution of single-stub problem.

every 1/2 wavelength. This is equivalent to going all the way around the Smith chart back to our original points.

Now that we know the appropriate position of the stub, all we have to do is to find its length. The length has to be such as to produce a normalized susceptance of either $-j1.6$ or $+j1.6$, depending on whether we connect to the first or second point that we found in Fig. 18-2. We can use the Smith chart to find the length of the stub (Fig. 18-3). The admittance of a shorted stub lies around the outer circle of the chart. On this path all of the admittances are pure susceptances because it is the locus of zero conductance. The admittance of a shorted stub starts at the bottom of the chart, where the conductance is infinite. Going around the chart toward the generator, we find that the normalized susceptance is $-j1.6$ at a distance of 0.089 wavelength. Thus the stub should be 0.089 wavelength long. If we were to place the stub at the second of the two points that we found in Fig. 18-2, we would need a normalized susceptance

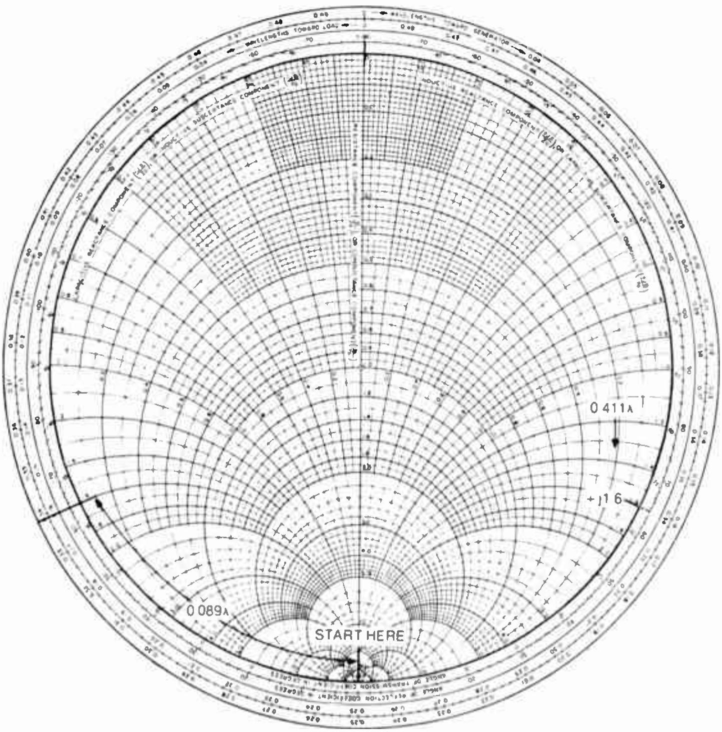


Fig. 18-3. Finding stub length.

of $+j1.6$. As shown in Fig. 18-3, this requires a stub length of 0.411 wavelength.

In the above example, we found two different places on the transmission line where we could place a shorted stub to accomplish the desired impedance transformation. Inasmuch as line losses increase with the length of the line, we usually choose the point closest to the load. The two solutions to the problem are shown in Fig. 18-4. Note that the impedance seen looking into the matching arrangement is 50 ohms, which is what we wanted. The standing-wave ratio on a 50-ohm line connected to this point will be 1:1. That is, there will be no reflection from this point. The standing-wave ratio on the line between the stub and the load will be a little over 4:1.

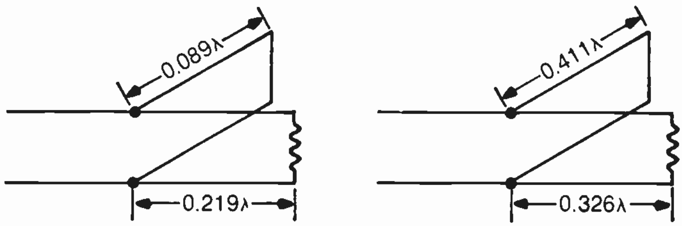


Fig. 18-4. Arrangements that satisfy the requirements in the single-stub problem.

The single-stub tuner is simple to design and only involves adding one element to the transmission line. The principal problem involves using it on a coaxial line. The shorted stub isn't difficult to manufacture; it consists of a section of coaxial line with a shorting plunger (Fig. 18-5). Unfortunately, the single-stub arrangement also requires an adjustable length of line from the load to the stub location. The manufacture of an arrangement that will provide an adjustable length of coaxial cable without disturbing its characteristic impedance is a nightmare. Sometimes the difficulty can be avoided by designing a matching section, then refining it by cut-and-try methods. Production models without adjustable sections can then be made from measurements. This has the disadvantage that adjustments cannot be made in the field.

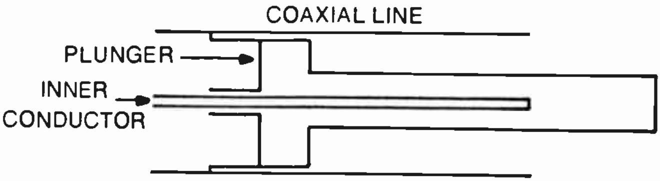


Fig. 18-5. Shorting stub.

Double Stubs

The difficulties of the single-stub arrangement can be avoided by using a double stub such as is shown in Fig. 18-6. Here we have two stubs placed across the line. They are separated by some distance other than 1/2 wavelength. Common separations are 1/8 and 3/8 wavelength. The design problem is to select the lengths of the shorted stubs so that the impedance seen at the second stub is equal to the characteristic impedance of the transmission line. The problem is solved easiest by the use of a Smith chart. The length l_1 of the first stub is chosen so that the admittance across the line at the point where the second stub will go will be on the unit-conductance circle of the Smith chart. The length l_2 of the second stub is then chosen to tune out the susceptance.

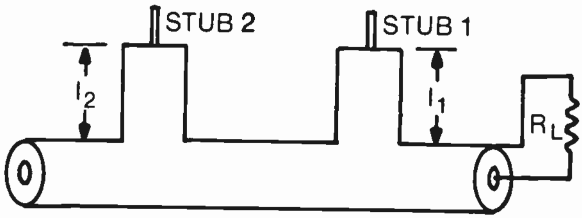


Fig. 18-6. Double stubs on coaxial line.

The big problem is to find the length of the first stub. The easy way is to select a stub spacing that makes it easy to use the Smith chart. When the spacing between the stubs is 1/8 wavelength, the solution is easy. We want the admittance at the place where the second stub is to be located to lie on the unit-conductance circle in Fig. 18-7 . Since the admittance at a point 1/8 wavelength away is shown on the Smith chart by rotating 45° toward the load, we want the admittance at the point where the first stub is located to fall on the circle marked

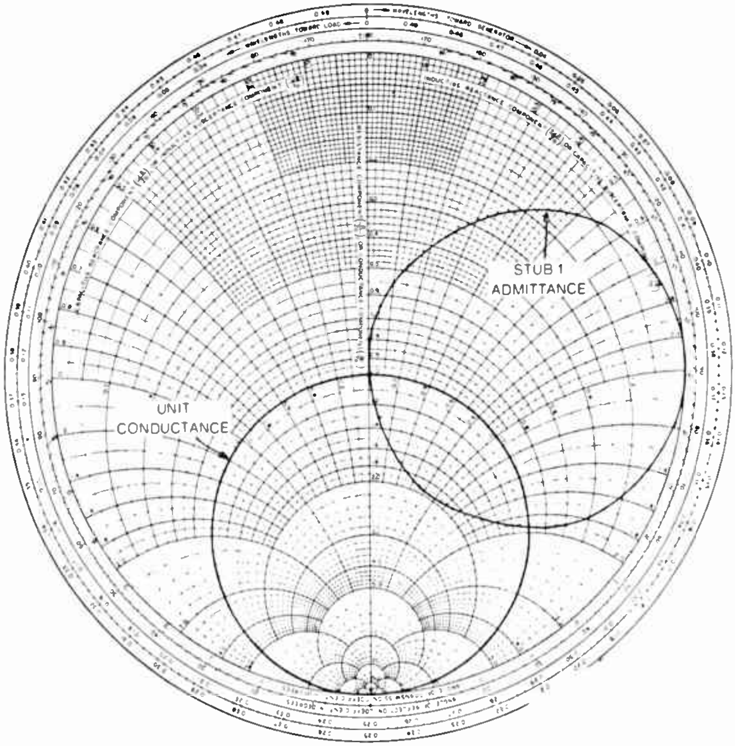


Fig. 18-7. Finding the circle for stub 1 admittance.

stub 1 admittance. All that we have to do is to adjust the length of the first stub so the admittance at that point will lie on the circle. Then we can be sure that the admittance at the location of the second stub, which is 1/8 wavelength toward the generator, will lie on the unit-conductance circle. The second stub will cancel the susceptance and the match will be perfect.

Double-Stub Problem

Let us consider the case where the load impedance is 100 + j100 ohms and the characteristic impedance of the transmission line is 50 ohms. We wish to match the load impedance to the characteristic impedance of the line, using the double-stub arrangement shown in Fig. 18-8. As in the single-stub example, we normalize the load impedance to 50 ohms and then find the normalized admittance. As in the preceding case, the normalized admittance is 0.25—j0.25 and

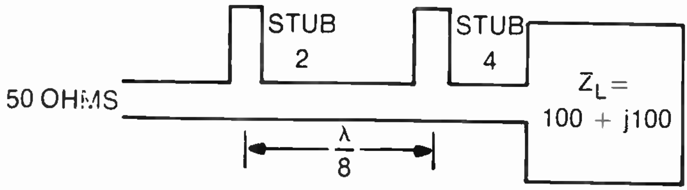


Fig. 18-8. Double-stub problem.

there are two solutions to the problem. Looking at the Smith chart in Fig. 18-9, we see that by adding a susceptance of $+j0.59$ to the load admittance, the admittance will become $0.25 + j0.34$, which lies on the circle where we want it. We could use a Smith chart to find that this corresponds to a stub length of 0.335 wavelength. Now we continue from this point 1/8 wavelength toward the generator and find the admittance $1 + j0.163$, which is on the unit-conductance circle, again, as wanted. Now we merely have to adjust the second stub length

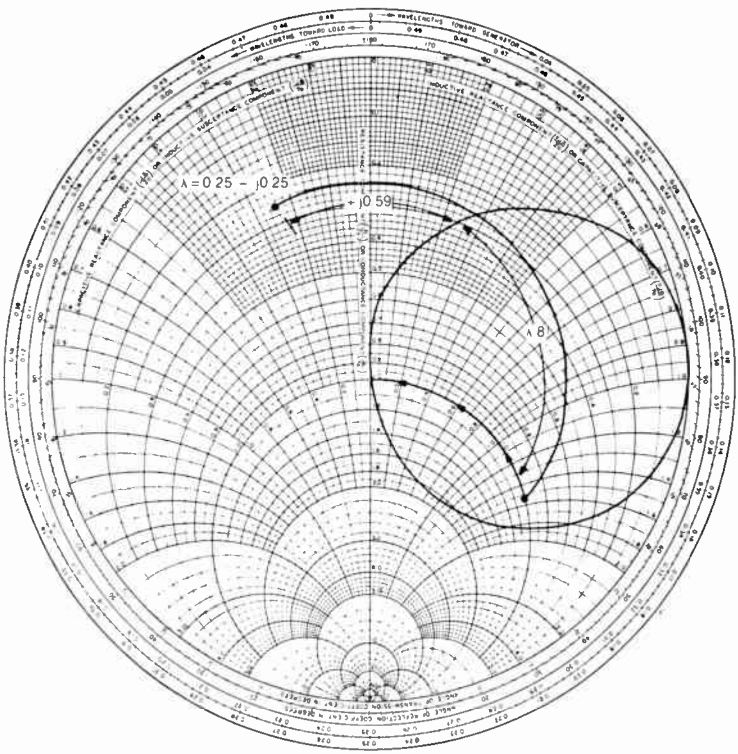


Fig. 18-9. Smith chart solution to double-stub problem.

so that its susceptance is $-j1.63$. From a Smith chart we would find that this corresponds to stub length of 0.088 wavelength. The final solution is shown in Fig. 18-10.

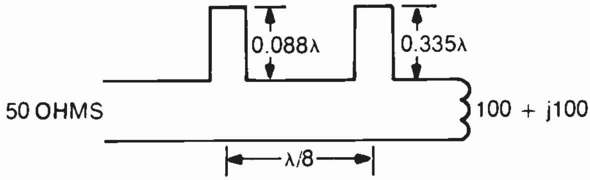


Fig. 18-10. Final answer to double-stub problem.

Not all possible values of load impedance can be matched with any given stub spacing. The widest possible range of load impedances can be matched when the spacing between stubs is 1/2 wavelength; however, at this spacing the adjustment of the lengths of the stubs is so critical as to be impractical. The smallest range of load impedances can be matched when the spacing between stubs is 1/4 wavelength. A spacing of 1/8 or 3/8 wavelength represents a good compromise between critical adjustment and the range of impedances that can be matched. If an impedance falls outside of the range covered by a particular stub spacing, it can be handled by adding a small fixed length of line between the load and the first stub.

Stub impedance-matching arrangements are often called *tuners*. The commonest types are the single-stub tuner and the double-stub tuner but it is possible to have more than two stubs. All arrangements operate on the same principles.

TAPERED LINE SECTION

A tapered line section that is a part of the antenna itself is sometimes used for impedance matching. Several different types of tapering have been used, including linear (A in Fig. 18-11) exponential (B), and hyperbolic. In all of them the dimensions of the line are slowly changed from a size that gives one value of characteristic impedance to one that gives another value. In all cases, the tapered section must be at least 1/4 wavelength long. This means that a tapered line section has a high-pass mode of operation. At all frequencies above a certain frequency, the tapered section will be a 1/4 wavelength or longer, and the impedance transformation will be

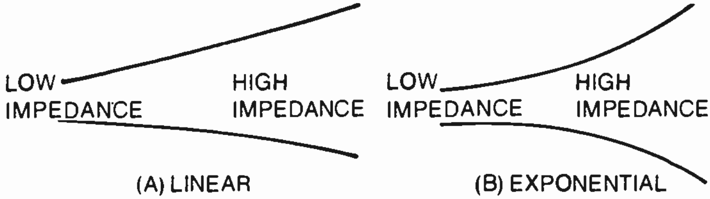


Fig. 18-11. Tapered line sections.

accomplished. At frequencies below this value, the length of the tapered section will be less than a 1/4 wavelength, and there will be a reflection.

QUARTER-WAVE TRANSFORMER

A quarter-wave transmission line has a driving-point impedance that is related to the load impedance by the equation

$$Z_{in} = \frac{Z_0^{\ 2}}{Z_L}$$

where Z_{in} = driving-point impedance
Z_L = load impedance
Z_0 = characteristic impedance of line

This means that by properly selecting the characteristic impedance of a quarter-wave section of transmission line, we can transform a load impedance to a desired value. Inasmuch as a 1/4 wavelength is equal to 90°, the quarter-wave transformer (Fig. 18-12) is very similar to the 90° *T*-network. It can match perfectly any two values of resistance. The quarter-wave transformer is frequency sensitive because the line section can only be a 1/4 wavelength long at one frequency. The bandwidth can be improved by using several such sections in tandem, with each one designed for optimum operation at a slightly different frequency.

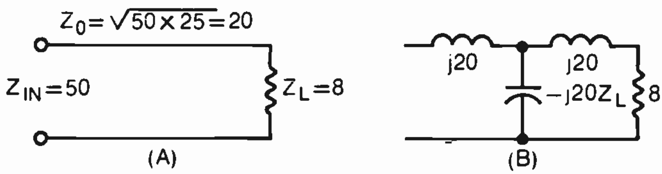


Fig. 18-12. Quarter-wave transformer with analogous T-network.

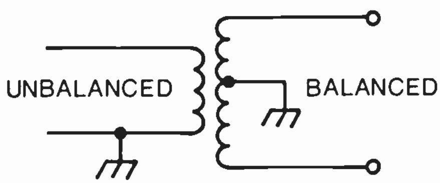
BALUN

Most FM and TV antennas are balanced; that is, they have two terminals, neither of which is at ground potential. The coaxial transmission lines, on the other hand, are unbalanced, in that the outer conductor is grounded. Thus we need some method of connecting a balanced load to an unbalanced transmission line. Devices that accomplish this are called *baluns*, from *bal*anced-to-*un*balanced transformer. If such a transformer is not used when driving a balanced load from a coaxial cable, the currents in the inner and outer conductors of the cable will not be equal, and the cable will radiate.

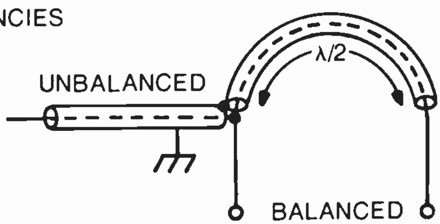
At low frequencies we can transform from an unbalanced source to a balanced load with the transformer shown in Fig. 18-13A. This arrangement is used to drive a push-pull stage from a single-ended stage. At higher frequencies such an arrangement would be very critical so a balun is used.

One type of balun that is easy to understand is shown in Fig. 18-13B. This device takes advantage of the fact that the voltage and current are the same at both ends of a half-wave section of transmission line, but the phase of the signal passing through the section is retarded by 180°. This amounts to a reversal in polarity.

There is an impedance transformation of 1:4 through the balun of Fig. 18-13B. This isn't easy to see at first, but in Fig. 8-14 we see that the voltage across the load is twice the voltage



(A) LOW FREQUENCIES



(B) RADIO FREQUENCIES

Fig. 18-13. Matching balanced loads to unbalance lines.

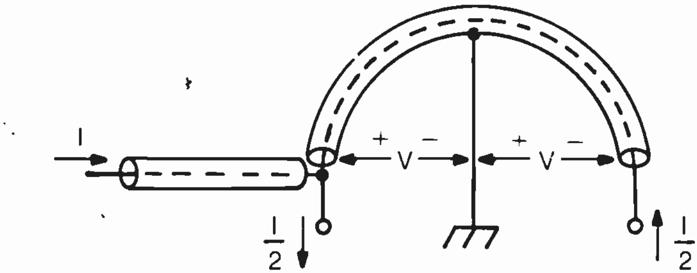


Fig. 18-14. One-to-four impedance transformation in a balun.

across the line, whereas the current in the load is one-half the current in the line. Dividing the load voltage by the load current gives us the impedance seen at this point.

$$\frac{2V}{I/2} = \frac{4V}{I}$$

where V/I is the input impedance.

Many different configurations of transmission-line elements are used for baluns. All have their advantages and limitations. In television proper operation over the required bandwidth is an important consideration.

HYBRID RING

A transmission-line element that can be used for many functions, including connecting two transmitters to a common load, is the hybrid ring (Fig. 18-15). This device is often made of coaxial components. It is sometimes called a *rat race*. The operation can be understood by tracing the paths through the ring, noting the phase shifts that are encountered through each path. The two paths from port 1 to port 4 are both the same length and have the same phase shift, so the signals traveling over the paths add at port 4. They also add at port 2. But the signal from port 1 cancels at port 3, and the signal from port 3 cancels at port 1.

AURAL AND VISUAL TV TRANSMISSION

A television station actually has two transmitters—one for the visual signal and one for the aural signal. There are several ways that these two signals can be radiated without interaction between the two transmitters. The simplest

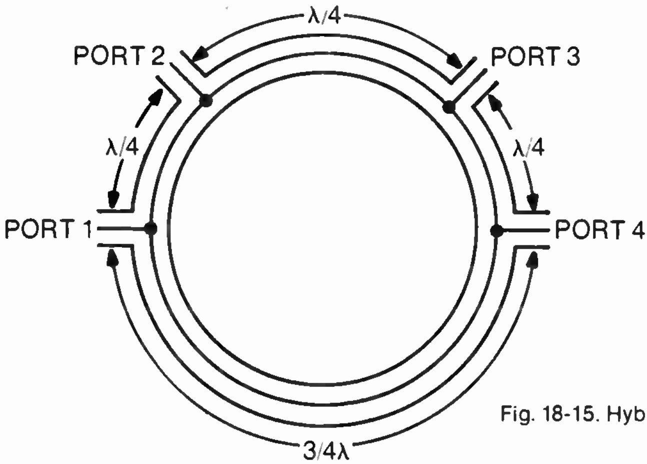


Fig. 18-15. Hybrid ring.

approach is to use two antennas. In this case, two separate transmission lines are run to the antennas.

When two separate antennas are used, the patterns must be as nearly alike as possible so that the sound will be heard wherever the picture is seen, and vice versa. Furthermore, the two antennas must be isolated from each other so that one does not pick up a substantial amount of the signal radiated by the other. Any interaction between the antennas would result in cross-modulation of the two signals. One solution is to use the top half of a superturnstile antenna to transmit the visual signal and the lower half to transmit the aural signal.

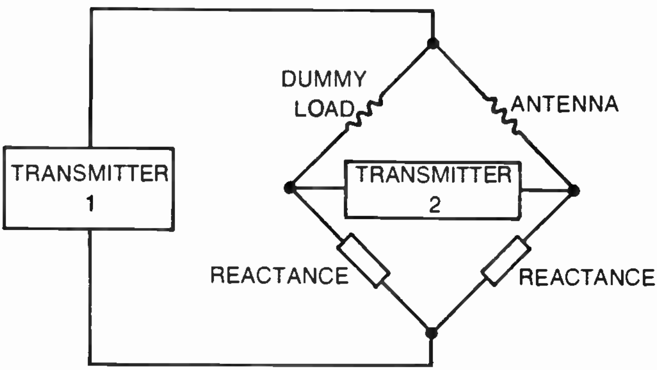


Fig. 18-16. Diplexer for connecting aural and visual TV transmitters to an antenna.

A common arrangement for combining the outputs of the visual and aural transmitters is a bridge circuit known as a diplexer (Fig. 18-16). This is used when there are two separate inputs to the antenna, one for north–south radiation and one for east–west radiation. The transmission line from the visual transmitter passes through a balun so that the signal is balanced to ground. The aural signal is fed from a coaxial line. The bridge is balanced so that no visual signal appears across the line from the aural transmitter and no aural signal appears across the line from the visual transmitter.

Bridge diplexers of this type are made in many different physical configurations that facilitate connecting them to antennas and transmission lines.

Chapter 19

FM and TV Antenna Measurements

There are completely different philosophies involved in the measurements made in FM or TV antenna systems and those made in AM antenna systems. In the AM antenna system the engineer is located within the confines of the antenna system most of the time. He walks between the elements (towers) making measurements and observations. He has access to the driving points of all of the elements. His measurements are made on the parts of the system, and he infers from them the quality of performance of the complete system.

In an FM or TV station the situation is completely different. The entire antenna is located on a tall tower, where only a steeplejack can gain access to the various elements. Impedance-matchiing networks as well as power-dividing networks are located up on the tower. The engineer must make most of his measurements on a single transmission line and from them determine whether the various elements of the system are operating properly.

Even the field-intensity measurements that are made to verify the coverage of the station are different. In the AM case the measurements are made along the surface of the earth, and the effect of the ground is determined by these measurements. An FM or TV field-intensity measurement should be made about 30 ft above the ground.

In making and interpreting measurements of an FM or TV antenna, the engineer is concerned with transmission-line theory. His measurements are made at one part of a line, and if they are not correct, he tries to infer what is wrong along the line or at the other end. Another basic difference between AM measurements and FM or TV measurements is in the instrumentation. In the AM station, instruments such as base-current meters and operating bridges are actually part of the antenna system. In the FM or TV antenna system there are usually no instruments at all. There is no point in having instruments at the separate elements, because they are located where the engineer cannot read them. The instrumentation used to make measurements on an FM or TV antenna is a part of the transmitter itself.

The transmitter is equipped with some type of transmission-line instrumentation that indicates voltage, current, or power, and usually, both forward and reflected parameters. For this reason, a working knowledge of transmission-line theory and the Smith chart is very handy for the engineer. It helps the engineer to use information that is available from instruments in the transmitter to diagnose troubles in the antenna system.

TRANSMISSION-LINE INSTRUMENTS

The transmission-line instruments commonly used on an FM or TV transmitter provide an indication of forward and reflected values of voltage, current, or power. There are several different arrangements that can be used for this purpose. One of the most common is the *directional coupler*. In its easiest-to-understand form, a directional coupler consists of an auxiliary transmission line loosely coupled to the main transmission line as shown in Fig. 19-1. The coupling from the main line causes waves to travel in the auxiliary line that are

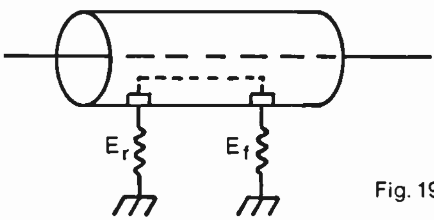


Fig. 19-1. Directional coupler.

proportional to the waves in the main line. Both ends of the auxiliary line are terminated in the characteristic impedance of the line so that there will be no reflection at either end. Thus the voltage across one terminating resistor is proportional to the forward voltage E_f in the main line, and the voltage across the other terminating resistor is proportional to the reflected voltage E_r in the main line.

A type of instrument called a *reflectometer* measures the ratio of these two voltages. This is a measure of the reflection coefficient K of the load impedance. If the magnitude of the reflection coefficient is known, it is an easy matter to find the standing-wave ratio on the line. It is given by

$$\text{VSWR} = \frac{1 + |K|}{1 - |K|}$$

where K is the magnitude of the reflection coefficient, and VSWR is the voltage standing-wave ratio.

FORWARD AND REFLECTED POWER

The terms *forward power* and *reflected power* cause a great deal of confusion. They should be clearly understood because their measurement is reasonably easy and tells a great deal about what is going on in a transmission line.

When the impedance of a load is not matched to the characteristic impedance of a transmission line, some of the energy that travels down the line toward the load is reflected back toward the transmitter. Power is the rate of flow of energy. Thus forward power is the rate at which energy propagates along the line toward the load. Reflected or reverse power is the rate at which energy is reflected back toward the transmitter. The power dissipated in, or radiated by, the load is equal to the difference between the forward and reflected power. This much seems reasonably clear. The two main questions are How much power is the transmitter delivering? and What happens to the reflected power?

Figure 19-2 represents a transmission line connected between a transmitter and an antenna, and in it is connected for measuring both forward and reverse power. The forward power is 1000W, and the reverse power is 110W. It is easy to see that the power radiated by the antenna is the difference, or

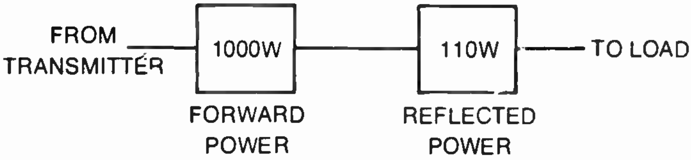


Fig. 19-2. Forward and reflected power in transmission line.

890W. The relationship between the standing-wave ratio on the line and the forward and reverse power is

$$\text{VSWR} = \frac{1 + \sqrt{\dfrac{P_r}{P_f}}}{1 - \sqrt{\dfrac{P_r}{P_f}}} = \frac{1 + \sqrt{\dfrac{110}{1000}}}{1 - \sqrt{\dfrac{100}{1000}}} = \frac{1.33}{0.67} = 2$$

If the characteristic impedance of the transmission line is 50 ohms, the load impedance connected to the receiving end of the line is either 25 or 100 ohms because

$$\text{VSWR} = \frac{R_L}{Z_0} \quad \text{or} \quad \frac{Z_0}{R_L}$$

Figure 19-3 is the same as Fig. 19-2 except that an impedance-matching network is added in Fig. 19-3. Now the transmitter sees a load impedance that is a pure resistance of 50 ohms. Since between the transmitter and the matching network the impedances are matched, there will be a standing-wave ratio of 1:1 on that section of the line. That is, there will be no reflection from the impedance-matching

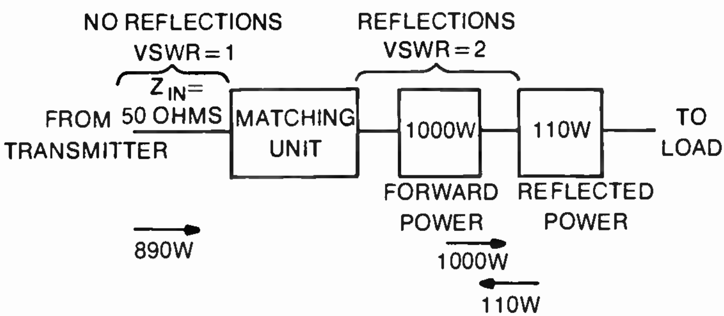


Fig. 19-3. Power from transmitter less than forward power in second section of line.

network. This, in turn, means that on the line between the transmitter and matching network the reflected power must be zero. The only power that we can measure on this section of the line is the power that goes to the right in the diagram and continues right out through the antenna. Thus the power in this section is all forward power and is equal to 890W. This shows that forward power and reflected power are merely ways of expressing what is happening on the transmission line. The fact that the forward power on the second section of the line is 1000W doesn't mean that the transmitter is putting out 1000W. It is merely a measure of the rate at which radiated energy plus energy stored on the line is moving toward the load. Likewise, the reverse power is merely a measure of the rate at which the energy stored in the line is moving toward the transmitter. Once this concept is clearly understood, forward power and reverse power show as clearly as anything what is going on in the transmission line.

ANALYSIS OF MEASUREMENTS

The purpose of making measurements on a feeder system is to verify that the system is operating properly, or to locate the trouble when it fails to operate properly. The biggest problem is to learn as much as possible about the operation of the antenna and feeder system by means of measurements made at one end of the transmission line. The Smith chart is a great help in interpreting whatever measurement data is available. It can be used to find various parameters of the system when other parameters are known.

When the forward voltage V_f and reflected voltage V_r on a line are known, we can draw a standing-wave circle on the chart as shown in Fig. 19-4, using the relationship

$$\text{VSWR} = 1 + \frac{V_r}{V_f} \bigg/ 1 - \frac{V_r}{V_f}$$

In the example of Fig. 19-4 the VSWR is 1.1:1.

If the distance in wavelengths to the load is known, we can use the chart to find the load impedance. This is the impedance at the end of the transmission line, which may be an antenna, but more often is a matching arrangement of some type. In most FM and TV stations the length of the line is several

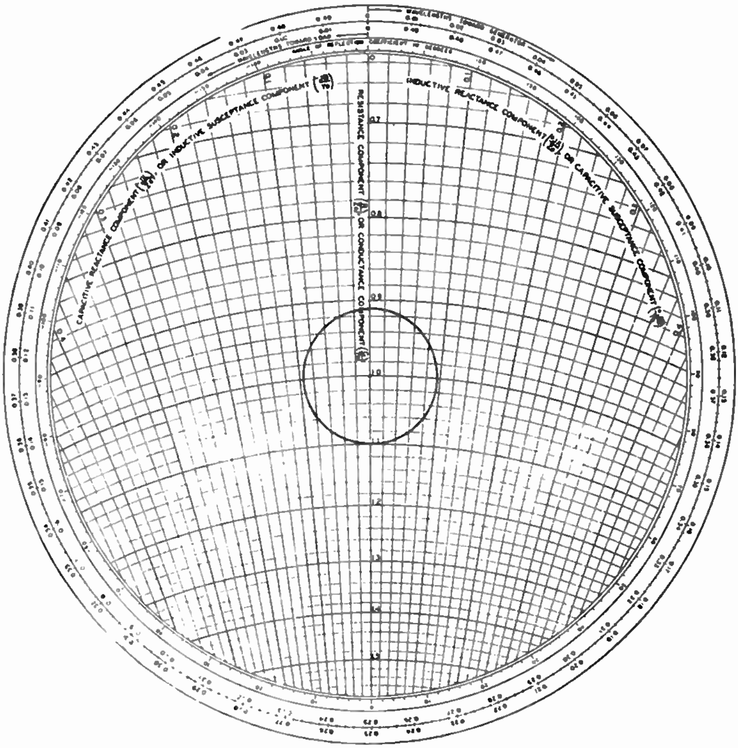


Fig. 19-4. Impedances on line having VSWR of 1:1.

wavelengths, so the standing-wave circle actually makes several turns around the Smith chart. All lines have some loss, so instead of the circle of Fig. 19-4, the path will actually spiral outward as we get closer to the load (Fig. 19-5). This tells us that with a normal line with no faults, the standing-wave ratio at the sending end will be less than it is at the load, because of the losses in the line.

Standing Waves as Fault Finders

A great deal of information about the system can be gained from merely knowing the standing-wave ratio at the sending end of the line. If the standing-wave ratio is very high, the line is probably open or shorted. Fortunately, nothing less than a catastrophic failure will cause the line to open, so we can assume that the fault is a short circuit. Open circuits on transmission lines can usually be located by visual inspection

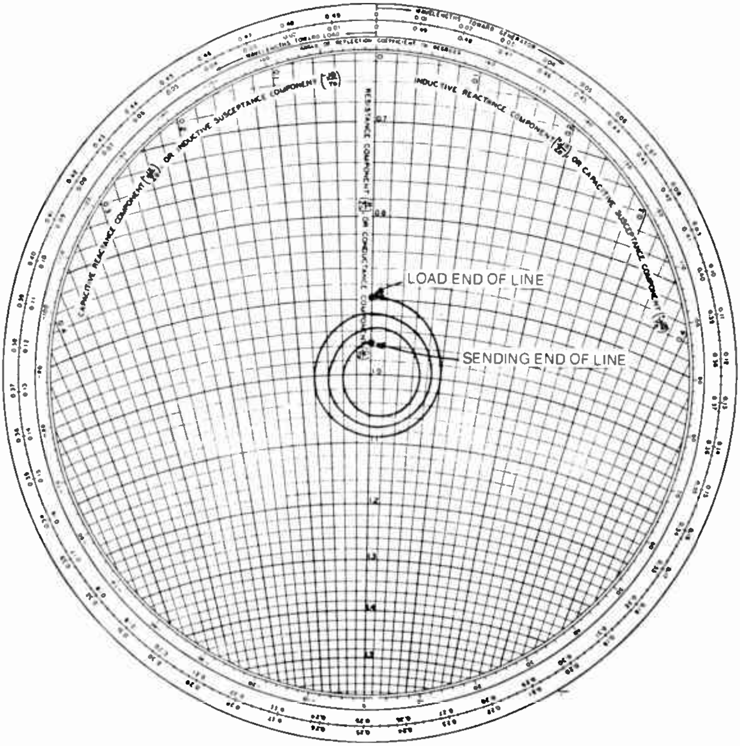


Fig. 19-5. Impedance on a line with losses.

with a pair of field glasses. A rough idea of where the short circuit is located along the line can sometimes be gleaned from the magnitude of the standing-wave ratio. If the normal losses of a line are known, we can calculate the maximum standing-wave ratio that will result from a short circuit at the load end of the line. This isn't conclusive, because the same standing-wave ratio that would result from a complete short circuit at the load might also result from a partial short circuit closer to the sending end.

An open circuit in one of the antenna elements can often be found from the standing-wave ratio. Suppose, for example, that we have the antenna shown in Fig. 19-6 which consists of six elements. Let us assume that when the system is operating normally, we have a standing-wave ratio of 1.1:1. When one of the elements opens, the standing-wave ratio will go up to about 1.2:1. (Fig. 19-6C). If additional elements open, the standing-wave ratio will go still higher.

Ghosting and Discontinuities

A fault in a TV feeder system can often be located by studying the ghost it causes on a TV receiver. The time required for one horizontal line of a TV picture is 53 μsec. During this time a signal can travel 26.150 ft from a point of reflection and back again. Thus we can use the distance by which a ghost trails the main picture on the screen of a TV set as a measure of the time between receiving the main picture element and the reflection. This time is a measure of the distance between the transmitter and the point on the line causing the reflection. The distance in feet is given by

$$\text{Distance} = 26.150 \frac{\text{ghost displacement in inches}}{\text{width of picture in inches}}$$

For example. if the ghost appears about 1/4 in. to the right of the picture on a TV screen that is 10 in. wide. the distance from the transmitter to the discontinuity that is causing the reflection is

$$\text{Distance} = 26{,}150 \frac{0.25}{10} = 65.4 \text{ ft}$$

This scheme is very helpful in locating faults along the line. When it is necessary to have someone climb towers to look for faults. the situation is simplified considerably if he knows approximately where to start looking. Although this

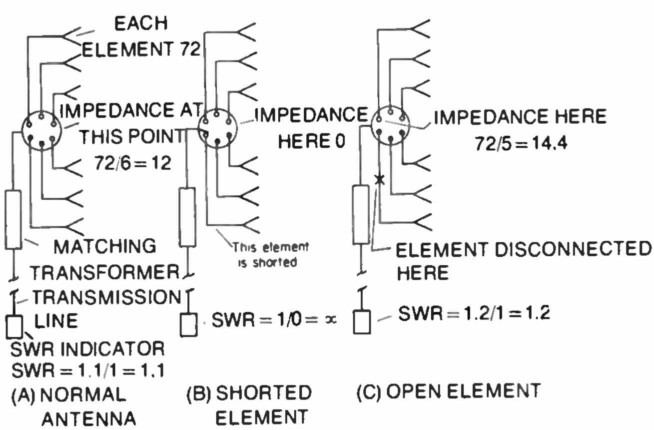


Fig. 19-6. Troubleshooting an antenna system.

principle can only be used with TV signals. there is a method of measurement called *time-domain reflectometry* that uses the same principle and can be used on any antenna system.

Time-Domain Reflectometry

A time-domain reflectometer is a device that transmits a pulse along a transmission line and displays the reflections on an oscilloscope. When the velocity factor of the transmission line is known. the horizontal axis of the oscilloscope display can be calibrated directly in terms of distance from the transmitter to the point where the reflection is being caused. Usually. in FM and TV systems the velocity of propagation on the transmission line is very close to the velocity of propagation in free space.

If a time-domain reflectometer is available. it provides an excellent way to keep track of slow changes in the feeder system. An oscilloscope camera can be used to record the reflection pattern when the system is operating properly. Then similar photographs can be taken at periodic intervals. Changes in the pattern indicate changes taking place in the system. A careful study of the patterns will show such changes in the system as slow corrosion of connections or water seepage into the system.

FIELD-INTENSITY MEASUREMENTS

Unlike a standard broadcast station. an FM or TV station is not required to make regular field-intensity measurements to assure that the radiating system is operating properly. In general. field-intensity measurements are only made when the station is originally put on the air or when major changes are made that influence the coverage of the station.

The field-intensity meter used for FM and TV measurements is similar in many ways to the instrument used for standard broadcast measurements. It consists of a calibrated receiver connected to an antenna. The antenna is usually a half-wave dipole rather than the loop type used in the standard broadcast meter. The dipole is preferably mounted on a mast with a rotator so that it can measure the field intensity at a height of up to 30 ft above the surface of the earth. Because of the frequencies involved in FM and TV broadcasting, the field intensity varies widely from one place to another on the surface of the earth. Measurements made at ground level would be meaningless; they would tell more about the environment of the measurement point than the field intensity of the signal.

Very frequently field-intensity measurements to establish the coverage of a station are made from a specially equipped vehicle. The vehicle usually has a telescoping mast that carries the antenna. Sometimes measurements are made with the vehicle in motion and the measurement being recorded on a strip chart recorder. Where possible, the vehicle travels on roads that are along radials from the antenna. This measurement is a highly specialized procedure and is usually performed by a consulting firm with the proper equipment.

Chapter 20
Lightning Protection

One of the most serious threats to the proper operation of any broadcast station is the effect of lightning or static charges. The antenna, the feeder system, and even the transmitter may be damaged or completely destroyed by a stroke of lightning. The problem of providing proper protection is complicated because many aspects of lightning are not very well understood, and the theories of lightning phenomena have the habit of changing every few years.

In addition to charges and lightning strokes associated with thunderstorms, static may build up on an antenna even when there is no lightning in the area. Usually, the buildup of an electric charge is greatest when there is a cool wind ahead of a rainstorm. The wind itself carries an electric charge, and this charge is imparted to conducting structures by the wind blowing over them. Perhaps the most striking demonstration of the effect of such static charges is the occasional arcing over of guy-wire insulators when there is no lightning in the area.

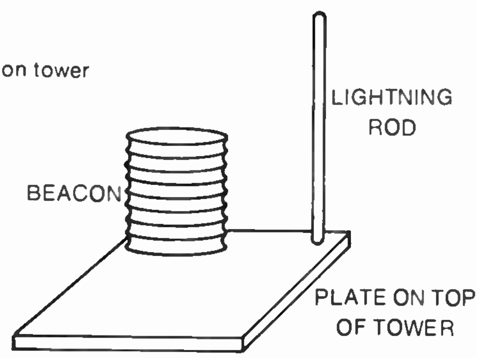
Although a lightning stroke is more damaging, a static buildup on a tower can cause serious problems. Every broadcast engineer knows that the old wives tale to the effect that lightning never strikes in the same place twice is just that—an old wives' tale. If lightning only struck once, most engineers would consider themselves fortunate indeed.

Lightning is a manifestation of static electricity —electricity that exists between two charged bodies. It is hard to predict how lightning will behave, because one of the charged bodies in this case probably isn't what we would ordinarily think of as a body at all. It is a cloud that is continuously changing in shape and location. Only its properties are very difficult to study. As a result of the uncertainty about lightning the conventional approach to lightning protection is to arrange things so that the damage that it does is minimized. A newer approach to the subject is based on avoiding lightning altogether.

DAMAGE PREVENTION

The first place to implement defensive tactics against lightning is at the top of the tower. A small lightning rod is usually attached to the top plate of the tower, beside the beacon (Fig. 20-1). The top of the rod is higher than the top of the beacon so that it will divert any lightning stroke away from the beacon, where it can destroy the lamp or fixture. This type of damage is not very extensive, but it is expensive to change a beacon at the top of a tower.

Fig 20-1. Lightning rod on tower



The guy wires should be protected by putting two or three insulators in series at the top of each guy wire where it fastens to the tower. This will cause any static-charge that accumulates on the guy wire to discharge into the ground rather than through the tower.

At the base of each tower an air gap is provided. In the event of a lightning stroke or static charge, the air in the gap ionizes and dissipates the charge before it gets into the feeder system. Two types of air gaps are commonly used. The *ball gap* of Fig. 20-2A can be spaced closer than other types because its smooth surface is less apt to support ionization from the transmitted signal. It is, however, subject to icing and should be provided with a rain shield.

The *horn gap*, shown in Fig. 20-2B, has the advantage of being *self-extinguishing*. When the gap ionizes because of a static discharge, the signal from the transmitter may sustain the arc. With the horn gap, the ionized air rises in the gap as it becomes heated by the arc. The distance between the electrodes of the gap increases with the height, so a point is soon reached where the gap is so large that the arc is extinguished.

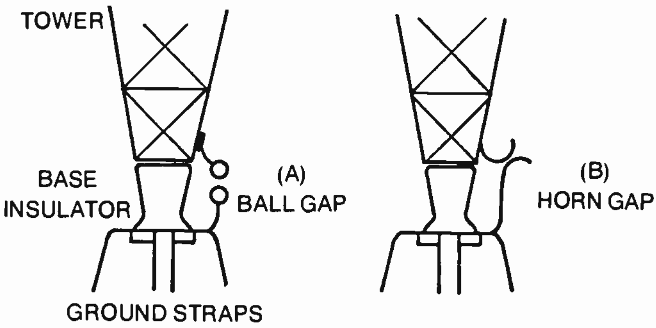


Fig. 20-2. Gaps for lightning protections.

To keep the static charge in the gap and out of the feeder system, many engineers put a 1- or 2-turn in the feeder going to the base of the tower (Fig. 20-3). The small inductance of the loop undoubtedly offers a high impedance to the steep waveform of a lightning discharge; however, if the turn should change shape, it will change the base impedance of the tower. This problem is more severe in tall towers that have a high base impedance and voltage.

Another device that is often used to prevent the buildup of static charges is the *static-drain choke*. This is a coil that has a very high impedance at the operating frequency, but a low DC resistance. It is connected to bleed off static charges on the tower before they have a chance to build up to damaging values. If a static-drain choke is used, it should be connected

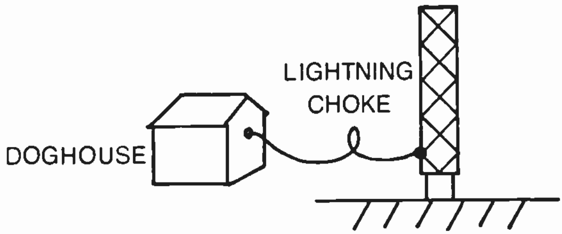


Fig. 20-3. Lightning choke.

on the line side of the base-impedance-matching unit, as shown in Fig. 20-4. This way the choke will be connected across the 50-ohm line impedance rather than across the base impedance of the tower. Base impedances cause enough problems of their own without the added uncertainty of a static-drain choke right across the base insulator. When a series capacitor is used in the impedance-matching network, the choke must be on the antenna side of the capacitor, but before the base-current meter.

The base-current meter is a favorite target for both lightning and static discharges. Many installations use a simple shorting switch for protecting the meter. Actually, this arrangement provides little protection, a lightning or static surge is of extremely short duration. Since power is the rate of flow of energy, a rapid surge causes extremely high voltages and currents. The resistance and inductance of the shorting switch, though small, will be high enough to cause damaging currents to flow through the meter. The best form of meter protection is the make-before-break switch, which removes the meter from the circuit when the meter is not in use.

ARC SUPPRESSION

One of the problems incidental to lightning and static discharges is that they can initiate an arc somewhere in the

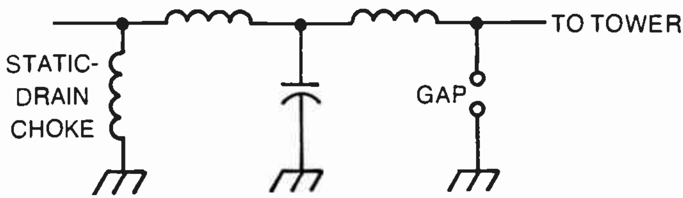


Fig. 20-4. Static-drain choke

system that is sustained by power from the transmitter. The initial discharge may not cause extensive damage, but the arc, if it is not extinguished, can cause serious damage. We hope that any arc set up by a static discharge will be across a spark gap, but this isn't always the case. The arc is sometimes inside a coaxial transmission line, where it can cause serious damage by melting the conductors or metallizing the spacing insulators. This situation is avoided by a circuit designed to automatically extinguish the arc (Fig. 20-5).

Somewhere between the arc-suppressing circuit and the antenna, there is a capacitor, so that the only DC path to ground is through the RF choke in Fig. 20-5. This path includes a relay and a source of operating current for the relay. Normally, no current flows through the relay, because there is no DC path. If an arc occurs, the current through the arc completes the circuit for the relay, thus energizing it. The contacts of the relay are connected to the transmitter in such a way as to remove power from the feeder system. Thus, whenever an arc occurs across the transmission line, the relay opens, removing the transmitter power long enough for the arc to dissipate.

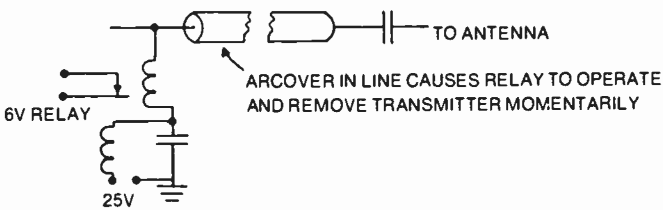


Fig. 20-5. Circuit to remove transmitter when arc occurs. (Courtesy Smith Electronics, Inc.)

LIGHTNING ELIMINATION

A comparatively new approach to the problem of lightning is *lightning elimination*. There has been a great deal of controversy about the merits of the system, but it has been in use long enough at many stations that its merits seem to be well established. The system is based on the principle of draining charges from the atmosphere before they can build up enough to cause a stroke of lightning. Many experts claim that it is not possible to drain off charges from clouds, and that the only function of a conductor such as a lightning rod is to

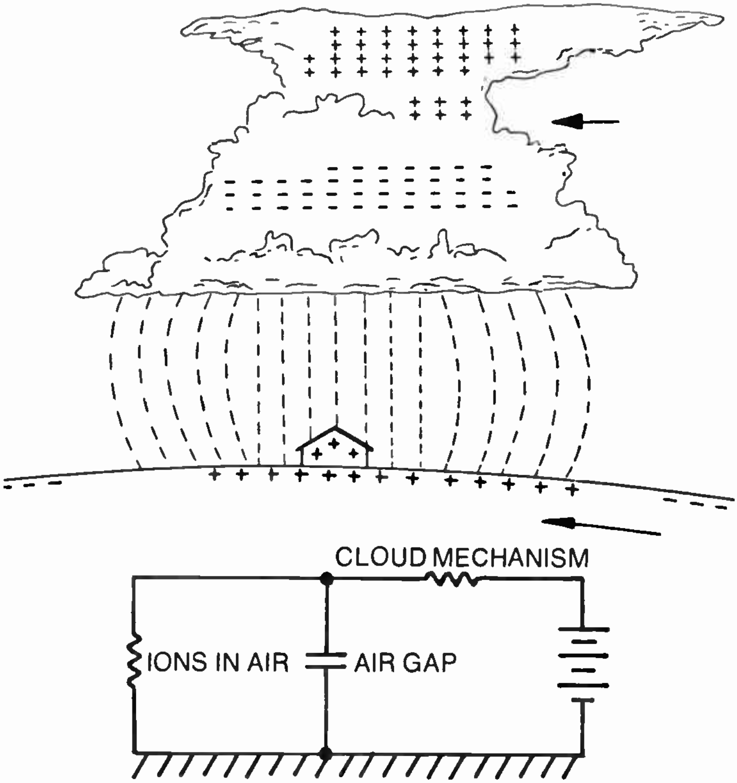


Fig. 20-6. Charges in thundercloud. (Courtesy of Smith Electronics, Inc.)

divert a stroke of lightning from objects to be protected. Probably, Ben Franklin, with the evidence of his kite experiment, would argue that it is indeed possible to drain charges from a cloud.

The new system operates on the principle that in a strong electric field, the air becomes ionized near a sharp point, and that the ions from the point migrate toward the opposite charge (Fig. 20-6). The *dissipation array* used in the new system actually has many thousands of sharp points, as many as 20,000. An array is placed on top of each tower and is connected to ground through a static-drain choke. When a strong field is present, ions form at the points and migrate toward the clouds, thus dissipating the charge. Stations using this system claim that damage from lightning has been reduced drastically.

Index